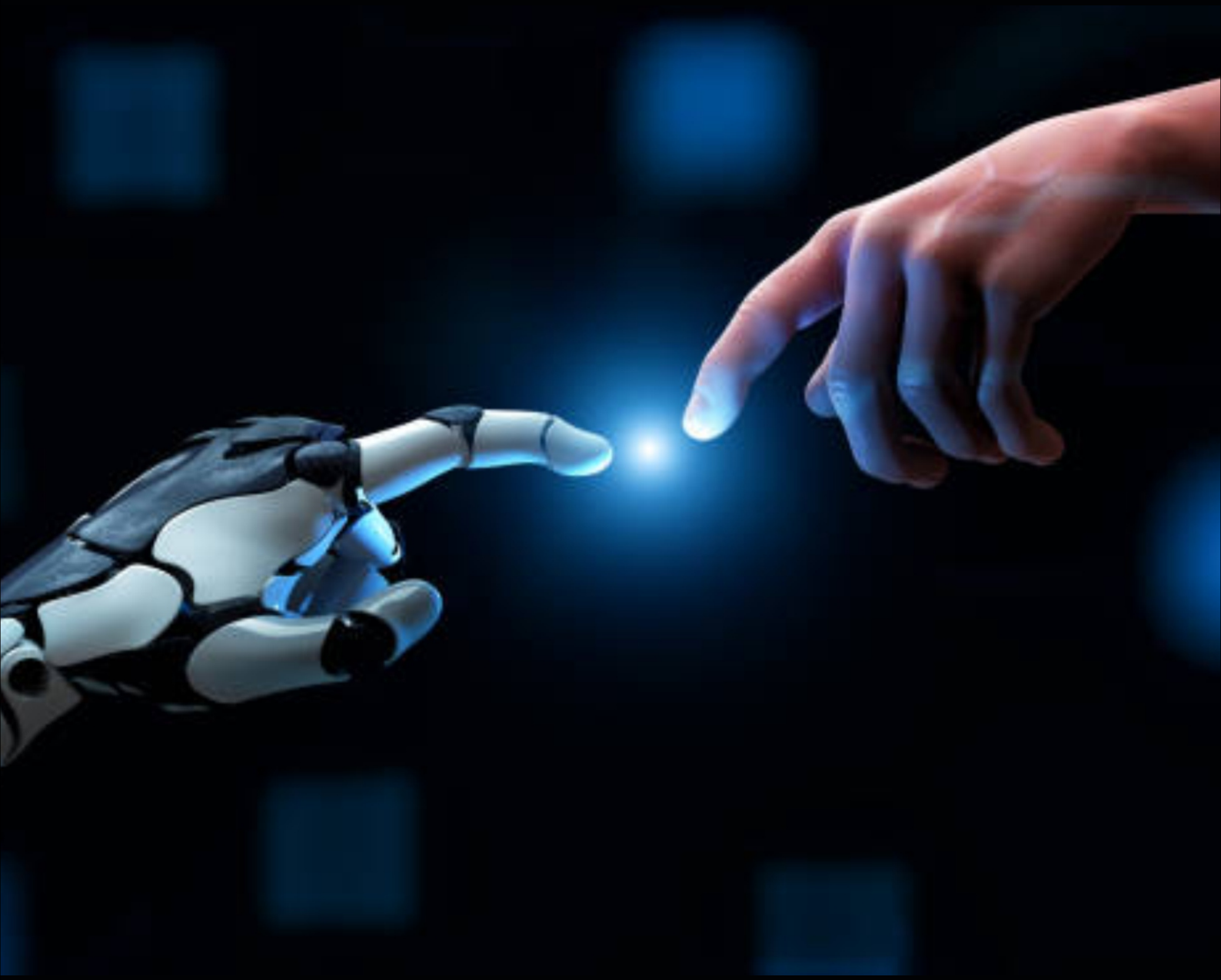
Dr. Joseph Teguh Santoso, S.Kom, M.Kom.

EVOLUSI AI OTONOM YANG BERWUJUD

YAYASAN PRIMA AGUS TEKNIK

Dr. Joseph Teguh Santoso, S.Kom, M.Kom.

EVOLUSI AI OTONOM YANG BERWUJUD



YAYASAN PRIMA AGUS TEKNIK

PENERBIT :
YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7695-42-0 (PDF)
9 786347 695420

EVOLUSI AI OTONOM YANG BERWUJUD

Penulis :
Dr. Joseph Teguh Santoso, S.Kom., M.Kom.

ISBN : 978-634-7695-42-0 (PDF)

Editor :
Dr. Agus Wibowo, M.Kom, M.Si, MM.

Penyunting :
Dr. Mars Caroline Wibowo. S.T., M.Mm.Tech

Desain Sampul dan Tata Letak :
Irdha Yunianto, S.Ds., M.Kom

Penebit :
Yayasan Prima Agus Teknik Bekerja sama dengan
Universitas Sains & Teknologi Komputer (Universitas STEKOM)

Anggota IKAPI No: 279 / ALB / JTE / 2023

Redaksi :
Jl. Majapahit no 605 Semarang
Telp. 08122925000
Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

Distributor Tunggal :
Universitas STEKOM
Jl. Majapahit no 605 Semarang
Telp. 08122925000
Fax. 024-6710144
Email : info@stekom.ac.id

Hak cipta dilindungi undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa ijin dari penulis

KATA PENGANTAR

Puji dan syukur senantiasa kita panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, hidayah, serta karunia-Nya, sehingga mahakarya pemikiran yang tertuang dalam buku "**Evolusi AI Otonom Yang Berwujud**" ini dapat diselesaikan dan hadir di hadapan para pembaca.

Selama bertahun-tahun, diskursus mengenai Kecerdasan Buatan (Artificial Intelligence) lebih banyak berkutat di ranah digital—algoritma yang terkurung di dalam peladen, memproses data, dan menghasilkan teks atau gambar. Namun, kita kini tengah berada di ambang revolusi teknologi yang jauh lebih fundamental. Kecerdasan tersebut mulai mengejawantah ke dalam bentuk fisik, berinteraksi langsung dengan hukum alam, dan menavigasi dunia nyata yang penuh dengan ketidakpastian. Buku ini lahir dari urgensi untuk mendokumentasikan transisi krusial tersebut: dari entitas perangkat lunak yang pasif, menjadi agen fisik otonom yang tidak hanya mampu belajar, tetapi juga berevolusi dan beradaptasi secara mandiri menghadapi tantangan lingkungan yang terus berubah. Mengingat kedalaman dan kompleksitas materi yang dibahas, buku ini dirancang secara terstruktur dan sistematis ke dalam enam bab utama. Setiap bab saling berkesinambungan membentuk pemahaman yang utuh:

Pada Bab 1: AI Berwujud, pembaca akan diletakkan pada fondasi dasar dari kecerdasan yang berinteraksi dengan dunia fisik. Bab ini tidak hanya menyajikan gambaran umum, tetapi juga membedah secara rinci teknologi dan komponen utama yang memungkinkan mesin untuk "merasakan" dan "bertindak". Kami juga mengulas arsitektur utama, aplikasi praktis, serta pentingnya simulator, kumpulan data (datasets), dan metrik penilaian yang menjadi laboratorium virtual bagi pengembangan AI jenis ini.

Selanjutnya, Bab 2: AI Berwujud Otonom membawa diskusi selangkah lebih maju. Memiliki tubuh fisik saja tidak cukup jika agen tersebut masih sangat bergantung pada kendali manusia. Bab ini mengeksplorasi teknologi dan komponen yang memungkinkan agen fisik beroperasi dengan tingkat otonomi yang tinggi. Pembaca akan memahami bagaimana mesin memproses informasi sensorik yang kompleks untuk mengambil keputusan secara seketika (real-time) di lingkungan yang dinamis.

Menyadari bahwa satu unit entitas memiliki batasan komputasi dan jangkauan spasial, Bab 3: AI Berwujud Kawanan Otonom hadir untuk memecahkan kebuntuan tersebut. Bab ini mengulas keterbatasan utama dari agen tunggal, lalu memperkenalkan konsep kecerdasan kolektif (swarm intelligence). Fokus pembahasan melingkupi bagaimana ribuan agen dapat saling berkolaborasi secara desentralisasi, membangun ketahanan operasional yang luar biasa saat terjadi kegagalan sistem, dan metrik apa saja yang digunakan untuk mengukur efektivitas pergerakan kawanan ini.

Puncak dari pergeseran paradigma dalam buku ini dimulai pada Bab 4: Kecerdasan Terwujud yang Berevolusi Sendiri. Di sinilah kita mendefinisikan ulang batas kemampuan mesin. Pembaca akan diajak menyelami teknologi dan komponen yang memungkinkan sistem

AI tidak hanya belajar dari data masa lalu, melainkan secara aktif memodifikasi struktur dan kognisinya sendiri agar selaras dengan lingkungan yang sama sekali baru, tanpa harus dilatih ulang dari awal oleh manusia.

Pemahaman tentang evolusi ini diperdalam secara sangat teknis pada Bab 5: Model Dunia yang Berevolusi Sendiri. Agar AI dapat bertahan dan berevolusi, ia memerlukan representasi internal atau "model dunia" mengenai bagaimana realitas ini bekerja. Bab ini mengupas tuntas landasan kognitif tersebut, mencakup mekanisme evolusi diri yang terjadi di dalam satu komponen sistem (intra-komponen) maupun yang melibatkan interaksi silang antar-komponen sistem secara adaptif.

Buku ini ditutup dengan manifestasi praktis pada Bab 6: Agen yang Berevolusi Sendiri. Kami menyatukan seluruh teori ke dalam tinjauan arsitektur dan strategi operasional agen mandiri. Guna memberikan gambaran yang lebih membumi, bab ini juga dilengkapi dengan studi kasus nyata yang memperlihatkan bagaimana agen-agen canggih ini mengimplementasikan konsep evolusi diri secara langsung di lapangan.

Karya ini dirancang tidak hanya sebagai buku teks akademis, tetapi juga sebagai peta jalan (roadmap) bagi para peneliti, insinyur robotika, mahasiswa, serta pembuat kebijakan yang ingin memahami arah masa depan teknologi kecerdasan buatan. Penulis menyadari sepenuhnya bahwa terwujudnya buku ini tidak lepas dari dukungan, diskusi kritis, dan kontribusi pemikiran dari banyak pihak. Untuk itu, penulis menyampaikan apresiasi yang setinggi-tingginya kepada segenap rekan sejawat, institusi penelitian, dan keluarga yang terus memberikan dorongan moral.

Evolusi kecerdasan buatan adalah sebuah perjalanan panjang yang belum mencapai titik akhir. Penulis berharap buku ini dapat menjadi pemantik gagasan baru, memperkaya literatur sains dan teknologi di Indonesia, serta menginspirasi lahirnya inovasi-inovasi brilian di masa depan. Kritik, saran, dan umpan balik yang konstruktif dari pembaca sangat kami nantikan demi penyempurnaan edisi-edisi berikutnya. Selamat membaca, dan mari bersama-sama menyongsong era kecerdasan yang berevolusi!

Semarang, Juli 2026
Penulis

Dr. Joseph Teguh Santoso, S.Kom., M.Kom.

DAFTAR ISI

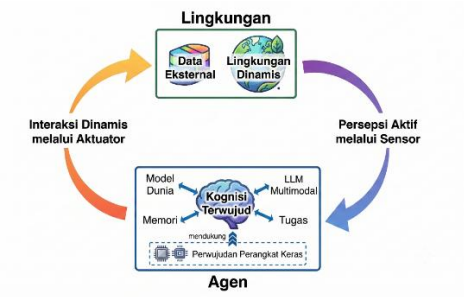
KATA PENGANTAR ..... **i**
DAFTAR ISI ..... **ii**
BAB 1 AI BERWUJUD **1**
1.1 Pendahuluan 1
1.2 Teknologi Utama AI Berwujud 3
1.3 Komponen Utama AI Berwujud 26
1.4 Arsitektur Utama AI Berwujud 43
1.5 Aplikasi Utama AI Berwujud 54
1.6 Simulator, Kumpulan Data, Dan Ukuran Penilaian 59
BAB 2 AI BERWUJUD OTONOM **70**
2.1 Pendahuluan 70
2.2 Teknologi Utama AI Berwujud Otonom 71
2.3 Komponen Utama AI Berwujud Otonom 78
BAB 3 AI BERWUJUD KAWANAN OTONOM **95**
3.1 Pendahuluan 95
3.2 Keterbatasan Utama Agen Tunggal Yang Terwujud 97
3.3 Teknologi Utama AI Berwujud Kawanan 103
3.4 Komponen Utama AI Berwujud Kawanan 114
3.5 Ketahanan Operasional AI Berwujud Kawanan 126
3.6 AI Berwujud Kawanan Otonom 129
3.7 Simulator, Dataset, Dan Metrik 139
BAB 4 KECERDASAN TERWUJUD YANG BEREVOLUSI SENDIRI **148**
4.1 Pendahuluan 148
4.2 Teknologi Utama Kecerdasan Terwujud Yang Berevolusi Sendiri 150
4.3 Komponen Utama Kecerdasan Terwujud Yang Berevolusi Sendiri 159
BAB 5 MODEL DUNIA YANG BEREVOLUSI SENDIRI **166**
5.1 Pendahuluan 166
5.2 Landasan Model Dunia Yang Berevolusi Sendiri 168
5.3 Evolusi Diri Di Dalam Komponen 173
5.4 Evolusi Diri Antar-Komponen 186
BAB 6 AGEN YANG BEREVOLUSI SENDIRI **193**
6.1 Pendahuluan 193
6.2 Komponen Utama Agen Yang Berevolusi Sendiri 194
6.3 Arsitektur Utama Agen Yang Berevolusi Sendiri 198
6.4 Strategi Utama Agen Yang Berevolusi Sendiri 202
6.5 Studi Kasus: Agen Yang Berevolusi Sendiri 206
DAFTAR PUSTAKA **212**

BAB 1
AI BERWUJUD

1.1 PENDAHULUAN

Ringkasan Kecerdasan Buatan (AI) Berwujud adalah cara menuju AI serbaguna (AGI), yang menjembatani dunia maya dan sistem fisik. Bab ini meninjau AI berwujud dari dasar hingga perkembangan terbaru, termasuk sejarah, teknologi utama, bagian, susunan sistem, aplikasi, program simulasi, kumpulan data, dan ukuran penilaian. Kemajuan terbaru dalam Model AI Bahasa Besar (LLM) dan model dunia telah sangat memajukan bidang ini. LLM meningkatkan AI berwujud melalui proses berpikir berdasarkan makna dan pemecahan tugas, memungkinkan instruksi tingkat tinggi dan tindakan bahasa alami. Model dunia menyediakan gambaran di dalam sistem dan kemampuan memperkirakan, mendukung interaksi yang sesuai dengan hukum fisika. Bab ini memperkenalkan pendekatan berdasarkan LLM/MLLM dan berdasarkan model dunia. Bab ini menjelaskan peran mereka dalam proses berpikir dan interaksi berwujud. Bab ini juga menyarankan susunan sistem gabungan yang menggabungkan MLLM dan model dunia untuk memungkinkan pelaksanaan tugas rumit di lingkungan fisik.

AI berwujud berasal dari Tes Turing Berwujud oleh Alan Turing pada tahun 1950, yang dirancang untuk mengeksplorasi apakah sistem AI dapat meniru kecerdasan manusia untuk mencapai AI serbaguna (AGI). Di antara mereka, sistem AI yang hanya memecahkan masalah yang tidak melibatkan dunia fisik di dunia digital (cyberspace) umumnya didefinisikan sebagai AI yang tidak berwujud, sedangkan mereka yang juga dapat berinteraksi dengan dunia fisik dianggap sebagai AI berwujud. AI berwujud (embodied AI) dibangun berdasarkan wawasan dasar dari ilmu tentang cara berpikir dan ilmu saraf, yang menyatakan bahwa kecerdasan muncul dari hubungan yang saling memengaruhi antara pengamatan, proses berpikir, dan interaksi. Seperti yang ditunjukkan pada Gambar 1.1, AI berwujud mencakup tiga bagian utama dalam suatu siklus tertutup, yaitu (1) Pengamatan Aktif (pengamatan lingkungan yang digerakkan oleh sensor), (2) Proses Berpikir Berwujud (pembaruan proses berpikir yang digerakkan oleh pengalaman sebelumnya), dan (3) Interaksi yang Menyesuaikan (kendali tindakan yang dilakukan melalui penggerak). **Selain** itu, Bentuk Perangkat Keras juga sangat penting karena meningkatnya kebutuhan perhitungan komputer dan energi, terutama di bawah batas waktu tunda dan daya perangkat dalam keadaan penggunaan dunia nyata.



Gambar 1.1 Konsep AI berwujud

Pengembangan AI berwujud telah berkembang dari cara yang memakai satu jenis data menjadi cara yang memakai banyak jenis data. Pada tahap awal, AI berwujud terutama dipelajari dengan berfokus pada bagian individual dengan jenis data tunggal, seperti penglihatan, bahasa, atau tindakan. Bagian pengamatan, proses berpikir, atau interaksi didorong oleh satu masukan dari sensor, misalnya, pengamatan cenderung didominasi oleh jenis data gambar, proses berpikir cenderung didominasi oleh jenis data bahasa, dan interaksi cenderung didominasi oleh jenis data tindakan. Meskipun metode ini berkinerja baik dalam bagian individual, metode ini dibatasi oleh jangkauan informasi yang sempit yang diberikan oleh setiap jenis data dan kesenjangan bawaan antara jenis data di seluruh bagian. Perkembangan berkelanjutan AI berwujud telah mengungkapkan keterbatasan pendekatan dengan satu jenis data, mendorong pergeseran besar menuju penggabungan berbagai jenis data dari sensor. Dengan demikian, AI berwujud dengan banyak jenis data secara alami muncul untuk menciptakan sistem AI yang lebih dapat menyesuaikan diri, luwes, dan tangguh yang mampu melakukan tugas-tugas rumit di lingkungan yang berubah-ubah. Model AI Bahasa Besar (LLM) memberdayakan AI berwujud melalui proses berpikir berdasarkan makna dan pemecahan tugas, membawa instruksi bahasa alami tingkat tinggi dan tindakan bahasa alami tingkat rendah ke dalam Proses Berpikir Berwujud.

Karya-karya representatif yang digerakkan oleh LLM termasuk SayCan, yang i) menyediakan kumpulan tindakan bahasa alami yang telah dilatih untuk dunia nyata untuk membatasi LLM agar tidak mengusulkan tindakan yang tidak layak dan tidak sesuai konteks; ii) menggunakan LLM untuk mengubah instruksi bahasa alami menjadi urutan tindakan bahasa alami; dan iii) memanfaatkan rumus nilai untuk memeriksa kelayakan urutan tindakan bahasa alami dalam lingkungan fisik tertentu. Karya-karya ini menunjukkan bahwa LLM sangat berguna bagi robot yang bertujuan untuk bertindak berdasarkan instruksi tingkat tinggi yang berlaku dalam jangka waktu tertentu yang diungkapkan dalam bahasa alami. Namun, LLM hanyalah bagian dari keseluruhan sistem AI berwujud (misalnya, Proses Berpikir Berwujud), yang dibatasi oleh kumpulan tindakan bahasa alami yang tetap dan lingkungan fisik tertentu, sehingga menyulitkan AI berwujud berdasarkan LLM untuk menyesuaikan diri dengan robot dan lingkungan baru.

Perkembangan terbaru dalam LLM dengan banyak jenis data (MLLM) dan model dunia telah membuka cakrawala baru dalam pengembangan AI berwujud. MLLM dapat bertindak pada keseluruhan sistem AI berwujud, menjembatani masukan dengan banyak jenis data tingkat tinggi dan urutan tindakan gerak tingkat rendah ke dalam penggunaan AI berwujud secara menyeluruh. Proses berpikir berdasarkan makna memanfaatkan pemahaman antarjenis data MLLM untuk menafsirkan makna dari masukan gambar, suara, atau sentuhan, misalnya, mengenali objek, menyimpulkan hubungan ruang, dan memperkirakan perubahan lingkungan. Bersamaan dengan itu, pemecahan tugas menggunakan logika berurutan MLLM untuk memecah tujuan rumit menjadi tugas kecil sambil terus menyesuaikan rencana berdasarkan umpan balik sensor. Namun, MLLM sering gagal untuk mengaitkan perkiraan

dengan perubahan yang sesuai dengan fisika dan menunjukkan penyesuaian langsung yang buruk terhadap umpan balik lingkungan.

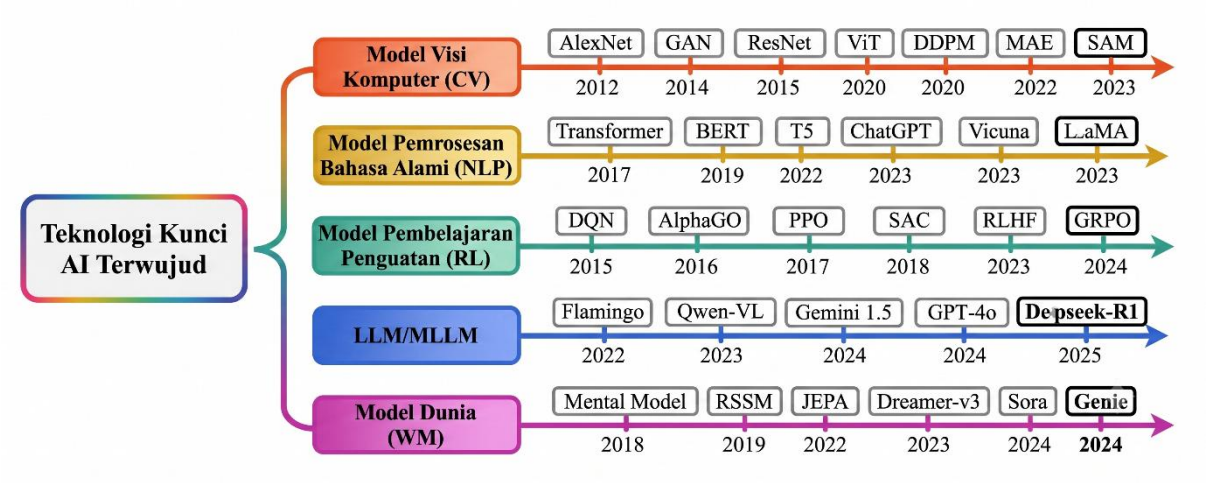
Di sisi lain, model dunia memberdayakan AI berwujud dengan membangun gambaran di dalam sistem dan membuat perkiraan keadaan mendatang dari dunia eksternal. AI berwujud berdasarkan model dunia semacam ini memudahkan interaksi yang sesuai dengan hukum fisika di lingkungan berubah-ubah. Gambaran di dalam sistem meringkas masukan dari sensor yang kaya ke dalam ruang data tersembunyi yang teratur, menangkap perubahan objek, hukum fisika, dan struktur ruang, serta memungkinkan sistem AI untuk berpikir tentang "apa yang ada" dan "bagaimana sesuatu berperilaku" di sekitarnya. Secara bersamaan, perkiraan keadaan mendatang menirukan potensi imbalan dari tindakan berurutan di berbagai rentang waktu yang selaras dengan hukum fisika, sehingga mencegah perilaku berisiko atau tidak efisien. Namun, pendekatan berdasarkan model dunia mengalami kesulitan dengan proses berpikir berdasarkan makna terbuka dan kekurangan kemampuan pemecahan tugas yang dapat diterapkan pada keadaan lain tanpa prioritas jelas.

Bab ini meninjau AI berwujud dari dasar hingga perkembangan terbaru, termasuk sejarah, teknologi utama, bagian, susunan sistem, aplikasi, program simulasi, kumpulan data, dan ukuran penilaian. Berdasarkan kemajuan di atas, kami selanjutnya berbagi wawasan kami tentang perlunya mengembangkan susunan sistem gabungan MLLM dan model dunia untuk AI berwujud, menekankan pentingnya dalam memungkinkan tugas-tugas rumit di dalam dunia fisik. MLLM memungkinkan proses berpikir tentang tugas sesuai keadaan tetapi mengabaikan batasan fisik, sementara model dunia unggul dalam simulasi yang mengikuti hukum fisika tetapi kurang memiliki pemahaman makna tingkat umum. Memanfaatkan MLLM dan model dunia secara bersamaan dapat menjembatani kemampuan memahami makna dengan interaksi fisik yang mendasar. Misalnya, EvoAgent mendesain sistem AI yang berkembang secara mandiri dengan susunan sistem gabungan MLLM dan model dunia, yang dapat secara mandiri menyelesaikan berbagai tugas jangka panjang di berbagai lingkungan melalui perencanaan mandiri, refleksi mandiri, dan pengendalian mandiri, tanpa campur tangan manusia. Kami percaya bahwa mendesain susunan sistem gabungan MLLM dan model dunia akan mendominasi sistem berwujud generasi berikutnya, menjembatani kesenjangan antara sistem AI khusus dan kecerdasan fisik umum.

1.2 TEKNOLOGI UTAMA AI BERWUJUD

AI berwujud (embodied Artificial Intelligence/embodied AI) berkaitan dengan pengembangan sistem AI yang beroperasi di dalam dan merespons kerumitan lingkungan simulasi atau fisik. Sistem-sistem AI ini harus terus-menerus merasakan, menafsirkan, dan bertindak terhadap lingkungan sekitarnya sambil menyesuaikan diri dengan perubahan yang terus terjadi dan umpan balik yang tidak pasti. Memungkinkan kemampuan tersebut tidak hanya membutuhkan cara kerja pengamatan dan kendali yang canggih, tetapi juga penggabungan yang lebih dalam antara pembelajaran, proses berpikir, dan interaksi. Seiring waktu, beberapa teknologi dasar telah terbukti sangat diperlukan dalam mendukung pengembangan embodied AI yang dapat menyesuaikan diri, tangguh, dan mampu diterapkan

pada keadaan lain. Bagian ini menyajikan lima teknologi utama seperti yang ditunjukkan pada Gambar 1.2. Teknologi-teknologi ini merupakan dasar cara kerja dari kecerdasan berwujud: Belajar melalui imbalan (Reinforcement Learning), Belajar terus-menerus (Continual Learning), Belajar dengan meniru (Imitation Learning), Model AI Bahasa Besar dengan Banyak Jenis Data (Dengan banyak jenis data Large Language Model), dan Model Dunia (World Model). Masing-masing bidang ini membahas dimensi yang berbeda dari proses pembelajaran berwujud.



Gambar 1.2 Kemajuan teknologi utama dalam AI berwujud

Belajar melalui imbalan menyediakan cara kerja baku untuk pengambilan keputusan berurutan melalui umpan balik lingkungan. Belajar terus-menerus membekali sistem AI dengan kemampuan untuk mengumpulkan dan mempertahankan pengetahuan dari waktu ke waktu. Belajar dengan meniru memudahkan perolehan pengetahuan dari perilaku ahli. Model AI Bahasa Besar dengan Banyak Jenis Data memungkinkan pemahaman dan interaksi tergabung di berbagai jenis data dari sensor. Model Dunia mendukung tiruan di dalam sistem dan proses berpikir jangka panjang. Teknologi-teknologi ini membentuk dasar yang selaras dan saling bergantung untuk membangun sistem AI yang mampu dan mandiri.

Belajar melalui imbalan

Belajar melalui imbalan (RL) menyediakan kerangka kerja dasar bagi sistem AI mandiri untuk memperoleh perilaku yang berfokus pada tujuan melalui interaksi coba-coba dengan lingkungannya. Tidak seperti pembelajaran terawasi, yang bergantung pada pasangan masukan-keluaran berlabel, RL dicirikan oleh tidak adanya instruksi jelas, melainkan menggunakan sinyal penghargaan berupa satu angka sebagai satu-satunya umpan balik. Pengaturan ini mencerminkan kondisi di mana organisme biologis sering belajar dan menyesuaikan diri, menjadikan RL sangat cocok untuk AI berwujud. Dalam cara ini, sistem AI berupaya mempelajari strategi yang memetakan pengamatan atau keadaan ke tindakan sehingga membuat imbalan sebesar mungkin total yang diharapkan dari waktu ke waktu, bahkan dalam menghadapi ketidakpastian, konsekuensi yang tertunda, dan perubahan yang rumit.

Secara baku, masalah RL sering dibingkai sebagai Proses Keputusan Markov (Markov Decision Process/MDP), yang didefinisikan oleh susunan nilai (S, A, P, R, γ), di mana S menunjukkan kumpulan keadaan, A kumpulan tindakan, $P(s' \mid s, a)$ peluang perpindahan antar keadaan, $R(s, a)$ aturan imbalan, dan $\gamma \in [0,1)$ pengurang nilai masa depan yang menyimpan dalam bentuk kode preferensi terhadap imbalan langsung dibandingkan imbalan masa depan. Pada setiap langkah waktu, sistem AI mengamati keadaan saat ini s_t, memilih tindakan a_t, menerima imbalan berupa satu angka r_t, dan berpindah ke keadaan baru s_{t+1}. Tujuannya adalah mempelajari strategi $\pi(a \mid s)$ yang membuat total imbalan sebesar mungkin yang diharapkan, yang biasanya didefinisikan sebagai jumlah setelah dikurangi nilainya dari imbalan masa depan $\mathbb{E}_\pi[\sum_t \gamma^t r_t]$.

Pengelompokan metode RL

Langkah kerja RL secara tradisional diklasifikasikan menjadi tiga kategori: metode berdasarkan nilai, metode berdasarkan strategi, dan metode aktor-penilai yang menggabungkan elemen keduanya. Setiap kategori mencerminkan cara kerja yang berbeda tentang bagaimana nilai dan strategi digambarkan atau diperbaiki. Subbagian berikut memberikan penjelasan yang lebih luas tentang setiap jenis. Metode Berdasarkan Nilai. Metode berdasarkan nilai bertujuan untuk mempelajari rumus nilai yang memperkirakan manfaat jangka panjang dari pengambilan tindakan tertentu dalam keadaan tertentu. Rumus nilai yang paling umum digunakan adalah rumus nilai tindakan $Q(s, a)$, yang mewakili imbalan total yang diharapkan mulai dari keadaan s, mengambil tindakan a, dan selanjutnya mengikuti suatu strategi. Setelah rumus nilai didekati, strategi dapat diturunkan secara tidak langsung dengan memilih tindakan yang membuat nilai sebesar mungkin yang diperkirakan.

Peran penting dalam kategori ini adalah Deep Q-Network (DQN), yang menunjukkan bahwa jaringan saraf konvolusional dapat digunakan untuk mempelajari fungsi Q secara langsung dari pengamatan gambar mentah. Untuk menstabilkan pelatihan, DQN memperkenalkan dua cara penting: jaringan target untuk mengurangi osilasi dan perbedaan dalam pembelajaran, dan pemutaran ulang pengalaman untuk memecah hubungan berdasarkan waktu dengan pengambilan sampel seragam dari tempat penyimpanan sementara. Sasaran belajar dalam DQN melibatkan membuat sekecil mungkin kesalahan selisih dari waktu ke waktu seperti pada Persamaan (1.1):

$$\mathcal{L}(\theta) = \left(Q_\theta(s_t, a_t) - \left[r_t + \gamma \max_{a'} Q_{\theta^-}(s_{t+1}, a')\right]\right)^2, \tag{1.1}$$

di mana θ dan θ^- adalah nilai pengatur dari jaringan saat ini dan jaringan target, masing-masing. Berdasarkan DQN, pengembangan lebih lanjut meningkatkan kinerja dan stabilitas. Double DQN mengatasi perkiraan yang terlalu tinggi dengan memisahkan langkah pemilihan dan penilaian tindakan. Dueling DQN memisahkan nilai suatu keadaan dari keuntungan tindakan tertentu menggunakan Persamaan (1.2).

$$Q(s, a) = V(s) + A(s, a). \tag{1.2}$$

Pemecahan ini meningkatkan kemampuan berlaku pada keadaan lain dengan memungkinkan model untuk menilai nilai keadaan secara independen dari tindakan spesifik. Metode berdasarkan nilai tetap sangat efektif dalam pilihan tindakan yang terpisah dan memiliki sedikit unsur. Namun, penerapannya pada tugas kendali berkelanjutan terbatas, karena persyaratan untuk menghitung maksimum di atas pilihan tindakan berkelanjutan seringkali membutuhkan banyak perhitungan komputer atau tidak dapat diatasi. Metode Berdasarkan Strategi. Metode berdasarkan strategi secara langsung memperbaiki strategi berparameter $\pi_\theta(a \mid s)$ dengan membuat total imbalan sebesar mungkin yang diharapkan, menawarkan alternatif mendasar terhadap pendekatan berdasarkan nilai. Metode ini sangat efektif dalam pilihan tindakan memiliki banyak unsur atau berkelanjutan, di mana penurunan strategi dari perkiraan nilai sering kali tidak praktis atau tidak stabil. Dasar teoretisnya terletak pada teorema arah perubahan strategi, yang memberikan ekspresi formal untuk menghitung arah perubahan pengembalian yang diharapkan terhadap nilai pengatur strategi seperti pada Persamaan (1.3).

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}[\nabla_\theta \log \pi_\theta\,(a \mid s) \cdot A^\pi(s,a)], \tag{1.3}$$

di mana $A^\pi(s,a)$ adalah fungsi keuntungan yang memperkirakan nilai relatif suatu tindakan dibandingkan dengan garis dasar yang diharapkan.

Langkah kerja REINFORCE merupakan aplikasi awal dari ide ini, menggunakan perkiraan Monte Carlo dari pengembalian dan memperbarui strategi melalui stochastic gradient ascent. Meskipun secara berdasarkan gagasan sederhana, REINFORCE mengalami tingkat perubahan yang tinggi dan kemampuan belajar dari sedikit data yang buruk. Perbaikan selanjutnya memperkenalkan batasan trust-region untuk memastikan pembaruan yang stabil. Trust Region Policy Optimization (TRPO) memperbaiki strategi di bawah batasan pada perbedaan KL antara strategi baru dan lama. Namun, ketergantungannya pada perbaikan tingkat kedua membuatnya kurang praktis. Untuk mengatasi hal ini, Proximal Policy Optimization (PPO) mengusulkan sasaran pengganti yang terpotong seperti pada Persamaan (1.4).

$$L^{\text{clip}}(\theta) = \mathbb{E}_t[\min(r_t A_t, \text{clip}(r_t, 1-\epsilon, 1+\epsilon)A_t)], \tag{1.4}$$

di mana r_t adalah rasio antara peluang strategi baru dan lama. Pembatasan yang sederhana namun efektif ini mencegah pembaruan yang terlalu agresif, menjadikan PPO sebagai dasar yang kuat untuk banyak patokan uji kendali berkelanjutan.

Metode berdasarkan strategi memungkinkan gambaran strategi yang luwes dan secara alami mendukung eksplorasi acak, sehingga sangat efektif dalam tugas kendali berkelanjutan dan memiliki banyak unsur. Metode ini tergabung dengan baik dengan susunan sistem belajar mendalam dan telah menunjukkan kinerja yang kuat dalam keadaan yang membutuhkan pengamatan dan tindakan ujung-ke-ujung. Namun, ketergantungan mereka pada perkiraan arah perubahan tingkat perubahan tinggi menyebabkan kemampuan belajar dari sedikit data yang buruk, terutama di lingkungan dengan imbalan yang jarang atau horizon yang panjang.

Keterbatasan ini telah mendorong pengembangan pendekatan hibrida, seperti metode aktor-penilai, yang berupaya menggabungkan keluwesan pembelajaran strategi dengan stabilitas perkiraan nilai.

Metode Aktor-penilai. Metode aktor-penilai menggabungkan kekuatan pendekatan berdasarkan nilai dan berdasarkan strategi dengan mempertahankan dua bagian terpisah: aktor yang mengusulkan tindakan dan penilai yang mengevaluasinya. Aktor memperbarui strategi ke arah yang disarankan oleh nilai perkiraan penilai, sementara penilai membuat sekecil mungkin kesalahan selisih dari waktu ke waktu untuk meningkatkan perkiraan nilai. Tujuan umum dalam metode aktor-penilai melibatkan perkiraan arah perubahan strategi menggunakan keunggulan penilai seperti pada Persamaan (1.5).

$$\nabla_\theta J(\theta) = \mathbb{E}_t \left[\nabla_\theta \log \pi_\theta \left(a_t \mid s_t \right) \cdot \left(r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t) \right) \right], \tag{1.5}$$

di mana V_ϕ adalah rumus nilai kondisi yang dipelajari dan diparameterisasi oleh ϕ.

Asynchronous Advantage Actor-Critic (A3C) merupakan tonggak sejarah yang memperkenalkan pelatihan paralel beberapa sistem AI di berbagai lingkungan. Hal ini dengan jelas meningkatkan jumlah contoh yang diproses dan stabilitas pembaruan. Dalam pengaturan kendali berkelanjutan, Deep Deterministic Policy Gradient (DDPG) menerapkan strategi pasti dengan pembelajaran di luar strategi yang sedang dipakai, menggunakan jaringan target terpisah untuk aktor dan penilai. Varian yang lebih baru dan lebih kuat, Soft Actor-Critic (SAC), menggabungkan pembatasan berdasarkan tingkat keacakan ke dalam tujuan untuk mendorong eksplorasi seperti Persamaan (1.6).

$$J(\pi) = \mathbb{E} \left[\sum_t r_t + \alpha \mathcal{H}\left(\pi(\cdot \mid s_t) \right) \right], \tag{1.6}$$

di mana $\mathcal{H}$ menunjukkan tingkat keacakan strategi dan α adalah pengatur tingkat keacakan yang menyeimbangkan eksplorasi dan perbaikan imbalan.

Metode aktor-penilai sangat cocok untuk masalah kendali berkelanjutan, di mana pembelajaran yang stabil dan gambaran strategi yang luwes sangat penting. Dengan menggabungkan perbaikan strategi (aktor) dengan perkiraan nilai (kritik), mereka memanfaatkan keunggulan dari kedua cara: aktor mengeksplorasi pilihan tindakan dan memperbarui strategi, sementara penilai memberikan perkiraan nilai tingkat perubahan rendah untuk memandu perbaikan ini. Struktur sinergis ini memungkinkan pembelajaran yang lebih efisien dalam hal sampel dan stabilitas yang lebih baik dibandingkan dengan pendekatan yang murni berdasarkan strategi atau nilai. Cara kerja aktor-penilai telah berhasil diterapkan dalam berbagai tugas berwujud, termasuk gerak berpindah, penggerakan benda, dan interaksi langsung, di mana sistem AI harus menyesuaikan diri dengan perubahan memiliki banyak unsur, berkelanjutan, dan diperpanjang secara berdasarkan waktu.

Perluasan dan Kemajuan dalam RL

Meskipun metode berdasarkan nilai, berdasarkan strategi, dan aktor-penilai membentuk inti dari RL modern, metode-metode ini sering kali menghadapi keterbatasan dalam hal kemampuan belajar dari sedikit data, koordinasi dalam keadaan banyak sistem AI, dan pembelajaran dalam rentang waktu yang panjang. Untuk mengatasi tantangan ini, beberapa perluasan telah diusulkan yang menggabungkan model terstruktur, perubahan interaksi sistem AI, dan kendali bertingkat. Kemajuan ini dengan jelas memperluas penggunaan RL dalam lingkungan rumit berwujud. Metode Belajar melalui imbalan inti yang diperkenalkan di atas telah mencapai keberhasilan yang luar biasa. Namun, metode-metode ini sering kali mengalami kemampuan belajar dari sedikit data yang buruk, terutama di lingkungan memiliki banyak unsur dengan imbalan yang jarang atau tertunda. Untuk mengatasi masalah ini, Belajar melalui imbalan berdasarkan model memperkenalkan model jelas perubahan lingkungan, yang memungkinkan sistem AI untuk menirukan lintasan masa depan dan belajar dari pengalaman yang dibayangkan daripada hanya mengandalkan interaksi dunia nyata.

Salah satu contoh paling awal, PILCO, menunjukkan penghematan data yang luar biasa dengan menggunakan model perubahan berdasarkan peluang dengan arah perubahan strategi analitik. Perkembangan selanjutnya, seperti World Models, menggabungkan perubahan tersembunyi yang menghasilkan data untuk mendukung perencanaan di lingkungan berdasarkan piksel. Keluarga langkah kerja Dreamer lebih lanjut memajukan arah ini dengan menggunakan model ruang keadaan berulang dan imajinasi tersembunyi, mencapai kinerja kompetitif dalam tugas kendali terpisah dan berkelanjutan dengan interaksi lingkungan yang jauh lebih sedikit. Karena tugas berwujud sering melibatkan banyak sistem AI yang berinteraksi, Belajar melalui imbalan telah diperluas ke pengaturan banyak sistem AI.

Belajar melalui imbalan dengan banyak sistem AI (MARL) memperkenalkan kerumitan tambahan karena keadaan yang terus berubah yang disebabkan oleh sistem AI pembelajaran lain, serta kebutuhan akan koordinasi atau kompetisi. Langkah kerja seperti Multi-Agent Deep Deterministic Policy Gradient (MADDPG), QMIX, dan Multi-Agent PPO (MAPPO) mengatasi tantangan ini melalui kerangka kerja pelatihan dipusatkan dan pelaksanaan dibagi ke tiap bagian, cara memilih bagian penting, dan strategi pemecahan nilai. Cara-cara ini telah diterapkan dalam berbagai bidang, mulai dari kendali robot bekerja sama hingga permainan strategis.

Belajar melalui imbalan bertingkat menghadirkan jalur riset lain yang berfokus pada penyederhanaan berdasarkan waktu. Dalam pengaturan ini, strategi tingkat tinggi memilih tujuan kecil atau tugas kecil, yang kemudian dilaksanakan oleh pengontrol tingkat rendah. Pemecahan ini memudahkan pembelajaran dalam lingkungan jangka panjang dan mendorong penggunaan kembali keterampilan. Karya perintis seperti FeUdal Networks dan HIRO telah menunjukkan efektivitas susunan sistem bertingkat dalam tugas-tugas dengan imbalan jarang dan banyak tahap.

Batasan RL

Dalam beberapa tahun terakhir, RL semakin tergabung dengan cara pembelajaran lainnya. Metode pembelajaran gambaran seperti CURL dan RAD meningkatkan efisiensi RL berdasarkan piksel dengan mempelajari bentuk angka ringkas yang relevan dengan tugas. Kerangka kerja belajar cara belajar seperti RL2 dan PEARL memungkinkan sistem AI untuk menyesuaikan diri dengan cepat terhadap lingkungan baru berdasarkan pengalaman sebelumnya. Tujuan yang diawasi sendiri, seperti yang ditunjukkan dalam UNREAL dan DrQ, semakin meningkatkan kemampuan belajar dari sedikit data dengan menambah RL dengan tugas-tugas tambahan.

Oleh karena itu, Belajar melalui imbalan telah berkembang menjadi bidang yang kaya dan beragam, menawarkan alat yang ampuh untuk pengambilan keputusan dalam ketidakpastian, perencanaan jangka panjang, dan kendali dapat menyesuaikan diri. Relevansinya terhadap kecerdasan berwujud sangat menonjol, karena secara alami menangkap interaksi siklus tertutup antara pengamatan, tindakan, dan imbalan yang mendefinisikan perilaku cerdas di lingkungan fisik. Dari permainan tindakan terpisah hingga kendali berkelanjutan sistem robotik, dari perbaikan sistem AI tunggal hingga koordinasi banyak sistem AI, dan dari pembelajaran tanpa model hingga imajinasi berdasarkan model, Belajar melalui imbalan terus memberikan dasar teoritis dan cara kerja praktis untuk pengembangan sistem AI mandiri serbaguna.

Seiring dengan meningkatnya kerumitan tantangan kecerdasan berwujud di dunia nyata, Belajar melalui imbalan akan tetap menjadi pusat untuk memungkinkan sistem AI belajar, menyesuaikan diri, dan berlaku pada keadaan lain di berbagai tugas, lingkungan, dan skala waktu. Integrasinya dengan kemajuan dalam pembuatan model dunia, pembelajaran gambaran, dan susunan proses berpikir kemungkinan akan memainkan peran kunci dalam membentuk generasi sistem mandiri berikutnya yang mampu beroperasi secara tangguh dalam lingkungan dunia nyata yang berubah-ubah, tidak pasti, dan tidak terstruktur.

Belajar terus-menerus

Dalam sistem AI berwujud di dunia nyata, sistem AI jarang ditempatkan di lingkungan tetap. Sebaliknya, mereka menghadapi serangkaian tugas, tujuan yang berubah, perubahan yang berubah, dan jangkauan waktu yang luas. Untuk tetap tangguh dan dapat menyesuaikan diri dalam lingkungan seperti itu, sistem AI harus belajar terus-menerus, mengumpulkan pengetahuan baru sambil mempertahankan kompetensi sebelumnya. Cara Belajar terus-menerus (CL) ini telah menjadi arah riset penting yang terletak di persimpangan pembelajaran mesin, ilmu tentang cara berpikir, dan robotika. Ini mencerminkan kebutuhan akan penyesuaian dan daya tahan seumur hidup di lingkungan yang berkembang dari waktu ke waktu dan tidak dapat sepenuhnya diantisipasi sebelumnya.

Metode CL bertujuan untuk mencapai dua tujuan mendasar: kemampuan menerima hal baru, kemampuan untuk menggabungkan pengetahuan baru secara efisien, dan stabilitas, kemampuan untuk mempertahankan informasi yang diperoleh sebelumnya tanpa gangguan. Menyeimbangkan kedua sifat ini sangat menantang karena fenomena yang dikenal sebagai lupa besar-besaran, di mana mempelajari tugas baru menurunkan kinerja pada tugas

sebelumnya. Masalah ini semakin diperkuat dalam Belajar melalui imbalan, di mana distribusi data terus berubah dan umpan balik yang tertunda membuat penggabungan pengetahuan menjadi lebih rumit.

Pada bagian ini, kami memeriksa pendekatan representatif untuk Belajar melalui imbalan berkelanjutan, termasuk metode berdasarkan pembatasan, cara kerja memori dan replay, dan kerangka kerja yang tidak bergantung pada tugas atau sadar konteks. Pendekatan-pendekatan ini menyediakan alat-alat penting untuk membangun sistem AI yang mampu menyesuaikan diri jangka panjang dalam lingkungan yang berubah-ubah dan terbuka.

Pengelompokan metode CL

Metode CL mengatasi tantangan dalam memperoleh dan mempertahankan pengetahuan di seluruh tugas berurutan atau distribusi data. Saat model terpapar informasi baru dari waktu ke waktu, mereka harus menyesuaikan diri tanpa mengorbankan kemampuan yang telah dipelajari sebelumnya. Untuk mendukung tujuan ini, berbagai metode telah dikembangkan, yang secara luas dapat diklasifikasikan menjadi tiga kategori: metode berdasarkan pembatasan, metode berdasarkan memori, dan strategi susunan sistem yang dapat berubah. Kategori-kategori ini mencerminkan asumsi yang berbeda tentang struktur tugas, aksesibilitas data, dan batasan perhitungan komputer. Pada bagian-bagian berikut, kami memeriksa cara-cara representatif dalam setiap kategori, menekankan cara kerja, kekuatan, dan keterbatasannya dalam konteks pembelajaran dapat menyesuaikan diri jangka panjang. Metode Berdasarkan Pembatasan. Metode berdasarkan pembatasan bertujuan untuk mengurangi lupa besar-besaran dengan membatasi pembaruan nilai pengatur yang paling mungkin mengganggu pengetahuan yang telah diperoleh sebelumnya. Metode-metode ini memperkenalkan istilah tambahan pada rumus kesalahan yang memberikan penalti terhadap perubahan nilai pengatur yang dikenali sebagai penting untuk tugas-tugas sebelumnya, sehingga mendorong stabilitas sekaligus memungkinkan kemampuan menerima hal baru yang berkelanjutan.

Salah satu pendekatan paling berpengaruh dalam kategori ini adalah Penguatan penyimpanan Bobot Elastis (EWC). EWC menggunakan Matriks Informasi Fisher untuk memperkirakan pentingnya setiap nilai pengatur berdasarkan kontribusinya terhadap kinerja tugas sebelumnya. Istilah penalti kuadrat ditambahkan ke sasaran perhitungan, yang mencegah penyimpangan dari nilai pengatur penting ini selama pembelajaran tugas baru. EWC telah berhasil diterapkan di berbagai bidang pembelajaran terawasi dan penguatan, termasuk patokan uji seperti Atari dan DeepMind Lab, dan telah menjadi dasar bagi banyak pengembangan selanjutnya. Beberapa ekstensi telah diusulkan untuk meningkatkan kemampuan berkembang dan kemampuan menyesuaikan diri EWC. Misalnya, Online EWC menggantikan perkiraan spesifik tugas tetap dengan aturan pembaruan online, sehingga dapat diterapkan pada keadaan aliran data langsung. Kecerdasan Sambungan saraf (SI) lebih lanjut meningkatkan penghematan perhitungan komputer dengan menghitung ukuran kepentingan tanpa meninjau kembali data sebelumnya selama pelatihan. Dalam lingkungan yang lebih rumit dan tidak pasti, khususnya dalam Belajar melalui imbalan, ekstensi Bayesian telah dieksplorasi. *Elastic Variational Continual Learning* (EVCL) menggabungkan penarikan

kesimpulan berbasis variasi dengan pembatasan dengan mempertahankan gambaran berdasarkan peluang atas nilai pengatur, memungkinkan sistem AI untuk mengelola perpindahan tugas secara lebih efektif.

Di luar pembuatan model berdasarkan peluang, cara kerja yang terinspirasi secara biologis menawarkan perspektif tambahan. *Memory Aware Synapses* (MAS) mengusulkan pendekatan pembatasan berdasarkan data berdasarkan penyesuaian agar selaras arah perubahan, yang memungkinkan perkiraan relevansi nilai pengatur tanpa tugas. Demikian pula, *Synaptic Intelligence* (SI) memperkenalkan rumusan yang dimotivasi oleh penguatan penyimpanan sambungan saraf dalam sistem biologis, di mana stabilitas koneksi yang dipelajari dimodulasi oleh perubahan pembelajaran. Pendekatan-pendekatan ini secara bersama-sama menunjukkan bahwa menggabungkan pentingnya tugas dan kestabilan sambungan saraf ke dalam tujuan perbaikan memberikan cara yang berprinsip dan efektif untuk mendukung Belajar terus-menerus baik dalam pengaturan buatan maupun berwujud. Metode berdasarkan pembatasan memberikan pendekatan yang berlandaskan teori dan efisien secara perhitungan komputer untuk Belajar terus-menerus. Telah ditunjukkan bahwa metode ini berkinerja baik dalam keadaan di mana batasan tugas diketahui atau dapat didekati, dan di mana penggunaan kembali nilai pengatur sangat penting. Namun, metode ini mungkin kurang efektif dalam pengaturan yang sangat tidak stasioner atau ketika informasi tugas tidak tersedia, sehingga memotivasi pengembangan cara pelengkap.

Metode Berdasarkan Replay. Metode berdasarkan replay mengatasi lupa besar-besaran dengan mempertahankan data masa lalu dan menyisipkannya dengan pengalaman baru selama pelatihan. Terinspirasi oleh sistem memori komplementer di otak manusia, khususnya interaksi hipokampus-neokorteks, pendekatan ini memungkinkan model untuk menguatkan penyimpanan pengetahuan jangka panjang sambil memperoleh informasi baru. Ide dasarnya adalah mendekati distribusi gabungan dari semua tugas yang diamati dengan mencampur sampel dari tugas saat ini dengan contoh representatif dari tugas sebelumnya. Penerapan umum melibatkan tempat penyimpanan sementara $\mathcal{M}$, yang menyimpan subset titik data yang telah dilihat sebelumnya. Ketika model menemukan data baru $\mathcal{D}_t$, model dilatih pada gabungan $\mathcal{D}_t \cup \mathcal{M}$, biasanya membuat sekecil mungkin nilai kesalahan total seperti Persamaan (1.7).

$$\mathcal{L}_{\text{total}} = \mathcal{L}(\mathcal{D}_t) + \lambda \mathcal{L}(\mathcal{M}), \tag{1.7}$$

di mana λ mengontrol kepentingan relatif dalam mempertahankan pengetahuan sebelumnya. Untuk mengelola kemampuan memori, cara pengambilan sampel seperti pengambilan sampel reservoir dan seleksi berdasarkan arah perubahan digunakan untuk memastikan keragaman dan relevansi tugas.

Experience Replay menunjukkan bahwa bahkan tempat penyimpanan sementara kecil pun dapat secara efektif mengurangi pelupaan, terutama jika dikombinasikan dengan pengambilan sampel strategis. Metode selanjutnya seperti *Multi-Timescale Replay* (MTR) meningkatkan daya simpan dengan menyimpan data pada beberapa skala berdasarkan waktu,

memungkinkan model untuk menangkap ketergantungan jangka pendek dan jangka panjang. Sistem hibrida juga telah muncul yang menggabungkan ingatan kejadian yang cepat berubah dengan memori makna yang lebih stabil untuk mempertahankan keluwesan dan stabilitas selama periode pembelajaran yang panjang. Ketika penyimpanan data mentah tidak memungkinkan karena masalah privasi, memori, atau hukum, pemutaran ulang dengan data buatan menawarkan alternatif. Dalam pendekatan ini, model yang menghasilkan data dilatih untuk mereproduksi distribusi data tugas sebelumnya. *Deep Generative Replay* (DGR) memperkenalkan ide ini dalam pengaturan klasifikasi, di mana pembuat data menghasilkan sampel buatan dari tugas-tugas masa lalu sementara pembelajar menyesuaikan diri dengan data baru. *Variational Continual Learning* (VCL) lebih lanjut menggabungkan pemutaran ulang dengan data buatan dengan jaringan saraf Bayesian, secara bersamaan mempelajari perkiraan setelah melihat data spesifik tugas dan menggunakan data sampel untuk meregularisasi pembaruan.

Meskipun metode berdasarkan replay termasuk di antara pendekatan yang paling banyak diadopsi dan secara empiris berhasil untuk Belajar terus-menerus, metode ini bukannya tanpa keterbatasan. Tantangan utama meliputi kemampuan memori yang terbatas, ketimpangan dalam pemilihan contoh, dan potensi penurunan kualitas yang menghasilkan data dari waktu ke waktu, terutama dalam horizon pelatihan yang panjang atau bidang memiliki banyak unsur. Metode Susunan sistem yang dapat berubah. Metode susunan sistem yang dapat berubah menangani Belajar terus-menerus dengan memodifikasi atau memperluas struktur model untuk mengakomodasi tugas, konteks, atau distribusi baru. Tidak seperti pendekatan berdasarkan pembatasan atau replay yang beroperasi di bawah susunan sistem tetap, metode ini memperkenalkan keluwesan pada tingkat susunan sistem untuk memisahkan gambaran tugas, mendorong daya simpan pengetahuan, dan memudahkan pembelajaran dapat menyesuaikan diri tanpa adanya batasan tugas yang jelas.

Salah satu pendekatan yang banyak dipelajari melibatkan penentuan jalur yang bergantung pada konteks, di mana aturan pemilih menentukan pengaktifan bagian kecil sesuai dengan masukan. Misalkan x menyatakan masukan dan $\{f_1, f_2, \dots, f_M\}$ adalah kumpulan bagian sistem fungsional. Aturan pemilih $g(x) \in [0,1]^M$ menetapkan bobot pada setiap bagian sistem, menghasilkan keluaran gabungan seperti pada Persamaan (1.8).

$$f(x) = \sum_{i=1}^{M} g_i\,(x) \cdot f_i(x). \tag{1.8}$$

Cara kerja ini memungkinkan model untuk mengkhususkan subset nilai pengatur yang berbeda untuk masukan yang berbeda, sehingga mengurangi interferensi antar tugas. Cara kerja Context-dependent Gating menggunakan gerbang biner yang telah ditentukan sebelumnya per tugas untuk menerapkan pengaktifan yang jarang, sementara ekstensi yang lebih baru menggunakan bentuk angka ringkas tentang keadaan yang dipelajari untuk memandu pemilihan bagian sistem secara tidak bergantung pada tugas. Strategi gating ini telah

menunjukkan peningkatan daya tahan dalam keadaan CL dengan secara efektif membagi kemampuan penggambaran.

Garis riset penting lainnya memanfaatkan jaringan pembuat nilai pengatur untuk menghasilkan nilai pengatur model secara berdasarkan keadaan berdasarkan bentuk angka ringkas tentang keadaan. Dalam pengaturan ini, jaringan pembuat nilai pengatur H memetakan vektor konteks $c \in \mathbb{R}^d$ ke nilai pengatur $\theta \in \mathbb{R}^n$ dari jaringan target f_θ. Rumusan ini memungkinkan jaringan target untuk menyesuaikan perilakunya sesuai dengan konteks tugas yang disimpulkan tanpa memerlukan model terpisah untuk setiap tugas. Metode berdasarkan jaringan pembuat nilai pengatur representatif menunjukkan bahwa pembangkitan nilai pengatur berdasarkan keadaan tersebut memungkinkan penyesuaian berkelanjutan sembari mempertahankan pengetahuan sebelumnya, bahkan tanpa batasan tugas jelas atau replay data.

Selain penentuan jalur selektif dan pengkondisian nilai pengatur, beberapa metode berfokus pada perluasan struktur jaringan dari waktu ke waktu. Kerangka kerja *Dynamic Expandable Network* (DEN) secara bertahap meningkatkan kemampuan model dengan memantau pergeseran penggambaran selama pelatihan. Ketika nilai pengatur yang ada tidak lagi cukup untuk menggambarkan data baru, neuron atau lapisan tambahan diperkenalkan. Pengetahuan yang diperoleh sebelumnya dipertahankan dengan membekukan nilai pengatur penting secara selektif, yang ditentukan melalui kriteria kepentingan berdasarkan kelangkaan. Strategi perluasan terkontrol ini menyediakan cara kerja untuk mengakomodasi informasi baru sambil mengurangi gangguan terhadap gambaran yang telah dipelajari sebelumnya.

Metode susunan sistem yang dapat berubah menawarkan dasar yang luwes dan mudah dikembangkan untuk Belajar terus-menerus, khususnya di lingkungan di mana batasan tugas ambigu atau sama sekali tidak ada. Dengan mendukung cara kerja seperti spesialisasi selektif, penyesuaian nilai pengatur berdasarkan keadaan, dan pertumbuhan struktural dapat menyesuaikan diri, pendekatan ini memungkinkan model untuk secara bertahap menggabungkan pengetahuan baru sambil membuat sekecil mungkin gangguan terhadap gambaran yang telah diperoleh sebelumnya. Kemampuan mereka untuk mengatur ulang sesuai perubahan jalur perhitungan komputer sebagai respons terhadap distribusi data yang berkembang sangat penting untuk membangun sistem AI yang mampu melakukan Belajar terus-menerus dan jangka panjang dalam pengaturan yang rumit dan terbuka di mana asumsi tradisional tentang struktur tugas mungkin tidak berlaku.

Perluasan dan Kemajuan dalam Belajar terus-menerus

Belajar terus-menerus telah berkembang melampaui cara dasarnya yaitu pembatasan, replay memori, dan penyesuaian susunan sistem. Seiring meningkatnya kerumitan lingkungan pembelajaran, rumusan yang lebih baru telah muncul untuk mengatasi tantangan yang rumit yang ditimbulkan oleh pengaturan dunia nyata, termasuk ketidakjelasan tugas, pergeseran bidang, dan penggabungan dari sensor dengan banyak jenis data. Salah satu arah penting melibatkan Belajar terus-menerus yang tidak bergantung pada tugas, di mana pembelajar tidak diberikan batasan tugas selama pelatihan. Alih-alih mengandalkan pembagian yang didefinisikan secara eksternal, model harus secara mandiri menyimpulkan struktur tugas atau

beroperasi secara efektif tanpa itu. Pendekatan yang terkenal termasuk penggunaan jaringan penarikan kesimpulan tugas, metode tanpa bentuk tetap Bayesian, dan cara pembelajaran gambaran dengan membandingkan yang mengelompokkan pengalaman berdasarkan struktur tersembunyi daripada identitas tugas. Cara-cara ini sangat relevan dalam konteks berwujud di mana batas tugas mungkin ambigu atau secara bawaan cair.

Ekstensi lain berpusat pada CL dengan banyak jenis data, di mana sistem AI harus menggabungkan dan mempertahankan pengetahuan di berbagai jenis data dari sensor, seperti penglihatan, bahasa, dan rasa posisi tubuh. Garis kerja ini mengeksplorasi penggabungan aliran masukan beragam dan dampaknya pada pelupaan dan transfer pengetahuan. Strategi seperti replay antarjenis data, gating spesifik jenis data, dan ruang bentuk angka ringkas yang selaras telah menunjukkan potensi dalam meningkatkan daya simpan dan kemampuan menyesuaikan diri dalam pengaturan dengan banyak sensor. CL juga telah diintegrasikan dengan belajar cara belajar, menghasilkan sistem yang tidak hanya mengumpulkan pengetahuan dari waktu ke waktu tetapi juga menjadi pembelajar yang lebih baik dalam prosesnya. Kerangka kerja meta-continual learning ini bertujuan untuk memperbaiki tidak hanya kinerja, tetapi juga perubahan pembelajaran yang mendasarinya, memungkinkan penyesuaian yang lebih cepat terhadap tugas-tugas baru dan kemampuan berlaku pada keadaan lain yang lebih baik. Metode seperti Meta-Experience Replay dan ANML mencontohkan tren ini dengan memanfaatkan ingatan kejadian bersamaan dengan aturan pembaruan dipelajari pada tingkat lanjut. Seiring bidang ini terus berkembang, batas-batas CL tidak hanya akan membutuhkan langkah kerja yang lebih kuat, tetapi juga wawasan teoritis yang lebih dalam dan kerangka penilaian yang lebih realistis. Kemampuan untuk belajar secara terus-menerus dan mandiri dalam lingkungan yang tidak terstruktur, dengan banyak jenis data, dan meluas secara berdasarkan waktu kemungkinan akan tetap menjadi kemampuan utama untuk kecerdasan berwujud tujuan umum.

Belajar dengan meniru

Dalam banyak keadaan AI berwujud, desain jelas fungsi penghargaan atau ketergantungan pada eksplorasi coba-coba yang ekstensif seringkali tidak layak, mahal, atau tidak aman. Belajar dengan meniru (IL), juga disebut sebagai pembelajaran dari contoh tindakan atau pembelajaran magang, menyediakan cara alternatif di mana sistem AI memperoleh strategi perilaku dengan mengamati contoh tindakan dari ahli. Pendekatan ini sangat berharga dalam konteks di mana keahlian manusia mudah tersedia, interaksi dunia nyata mahal, atau keselamatan saat digunakan menjadi perhatian utama. Dalam robotika dan sistem berwujud, Belajar dengan meniru membantu sistem AI mempelajari keterampilan dengan cepat. Ini menanamkan pengetahuan terstruktur melalui contoh tindakan dan mengurangi kebutuhan akan eksplorasi coba-coba.

Pengelompokan metode IL

Bagian ini meninjau metode dasar dan masa kini dalam Belajar dengan meniru. Kami menyusun bidang ini di bawah tiga tema utama: Penyalinan Perilaku, peniruan yang bersifat saling bersaing, dan metode peniruan bertingkat. Setiap kategori mewakili pendekatan pembuatan model yang berbeda, mencerminkan berbagai asumsi tentang ketersediaan data,

pengawasan, dan kebutuhan kemampuan berlaku pada keadaan lain. Penyalinan Perilaku memperlakukan peniruan sebagai masalah pembelajaran terawasi, memberikan dasar yang sederhana namun banyak digunakan. Metode saling bersaing mengambil inspirasi dari pembuatan model yang menghasilkan data dan Belajar melalui imbalan untuk mencocokkan perilaku ahli tanpa secara jelas membuat model imbalan. Metode bertingkat dan meta-peniruan bertujuan untuk meningkatkan keluwesan dan penghematan data, memungkinkan sistem AI untuk memperoleh keterampilan yang dapat ditransfer dan menyesuaikan diri dengan tugas-tugas baru dari contoh tindakan yang terbatas. Bagian-bagian berikut menyajikan ide-ide kunci, langkah kerja representatif, dan wawasan dalam setiap arah terkait cara kerja ini.

Penyalinan Perilaku. Penyalinan Perilaku (BC) adalah pendekatan paling mudah untuk Belajar dengan meniru, merumuskan masalah sebagai tugas pembelajaran terawasi yang memetakan pengamatan atau keadaan secara langsung ke tindakan ahli. Salah satu penerapan sukses paling awal adalah sistem ALVINN, yang menggunakan jaringan saraf untuk mengontrol kemudi berdasarkan gambar kamera yang menghadap ke depan untuk mengikuti jalan secara mandiri. Karya awal ini menunjukkan kelayakan mempelajari strategi gabungan penglihatan dan gerakan dari contoh tindakan. Sebuah studi baru-baru ini memperluas ide ini menggunakan jaringan konvolusional dalam untuk melakukan pengemudian ujung-ke-ujung di lingkungan simulasi dan dunia nyata, menetapkan patokan uji modern untuk kendali berdasarkan peniruan pada kendaraan mandiri.

Meskipun menarik dalam kesederhanaan dan kemudahan implementasinya, BC menderita keterbatasan kritis yang dikenal sebagai kesalahan yang berlipat ganda. Karena strategi hanya dilatih pada lintasan ahli, ia berkinerja buruk ketika menghadapi keadaan di luar distribusi selama penggunaan. Penyimpangan kecil dapat dengan cepat terakumulasi, menyebabkan penurunan kinerja, terutama dalam tugas-tugas jangka panjang atau yang kritis terhadap keselamatan. Untuk mengurangi keterbatasan ini, Belajar dengan mencari aturan imbalan (Inverse Reinforcement Learning/IRL) bertujuan untuk menyimpulkan aturan imbalan yang mendasari yang menjelaskan perilaku ahli. Sistem AI kemudian memperbaiki strategi menggunakan Belajar melalui imbalan di bawah imbalan yang disimpulkan ini. Karya penting memperkenalkan cara ini melalui pembelajaran magang, menunjukkan kemampuan berlaku pada keadaan lain yang lebih baik dibandingkan dengan BC. Selanjutnya, Maximum Entropy IRL telah memajukan metode perkiraan aturan imbalan untuk lingkungan yang rumit dan memiliki banyak unsur. Survei menyeluruh memberikan tinjauan sistematis tentang pendekatan-pendekatan ini. Survei tersebut menekankan kekuatan dan keterbatasan BC dan IRL, mengkontekstualisasikannya dalam cara kerja berdasarkan peluang dan cara memilih tindakan yang lebih luas.

Imitation Learning Saling bersaing. IL Saling bersaing memanfaatkan cara kerja pembuat data untuk melewati pembuatan model imbalan jelas. Peran penting adalah Generative Saling bersaing Imitation Learning (GAIL), yang merumuskan peniruan sebagai permainan dua pemain antara strategi (pembuat data) dan pembeda. Pembeda dilatih untuk membedakan lintasan ahli dari lintasan sistem AI, sementara strategi belajar untuk menipunya,

menghasilkan perilaku yang tidak dapat dibedakan dari ahli. Pengaturan saling bersaing ini memungkinkan sistem AI untuk secara tidak langsung menangkap maksud ahli tanpa memulihkan aturan imbalan jelas dan telah menunjukkan kinerja yang kuat dalam patokan uji kendali berkelanjutan seperti MuJoCo.

Namun, proses pelatihan saling bersaing GAIL diketahui sensitif terhadap pengaturan awal dan tidak efisien dalam pengambilan sampel. Untuk mengurangi kekurangan ini, Soft Q Imitation Learning (SQIL) mengubah peniruan menjadi RL di luar strategi yang sedang dipakai dengan memberikan imbalan positif tetap untuk perpindahan ahli dan nol untuk yang lain. Ini menghilangkan kebutuhan akan pembeda dan memungkinkan penggunaan langkah kerja di luar strategi yang sedang dipakai standar seperti DDPG atau SAC, meningkatkan stabilitas dan efisiensi pelatihan sambil mempertahankan kinerja yang kompetitif. Karya terbaru juga telah mengeksplorasi penstrukturan contoh tindakan untuk meningkatkan kinerja peniruan. Automatic Titik antara Extraction (AWE) membagi lintasan ahli menjadi titik antara tingkat tinggi, memungkinkan sistem AI untuk mempelajari strategi yang berpikir atas tujuan kecil abstrak. AWE meningkatkan penghematan data dan kemampuan berlaku pada keadaan lain, terutama dalam tugas robotika jangka panjang di mana peniruan tingkat rendah mungkin rapuh.

IL Bertingkat. IL Bertingkat mengatasi tantangan penskalaan peniruan ke tugas-tugas rumit dan berjangka panjang dengan memperkenalkan penyederhanaan berdasarkan waktu dan struktural ke dalam pembelajaran strategi. Alih-alih mempelajari strategi satu bagian utuh yang memetakan observasi langsung ke tindakan, kerangka kerja bertingkat memecah perilaku menjadi beberapa tingkat kendali. Strategi tingkat tinggi memilih di antara sub-strategi atau primitif perilaku berdasarkan konteks tugas, sementara setiap strategi tingkat rendah mengatur pelaksanaan motorik halus selama langkah waktu yang panjang. Struktur ini memungkinkan penggunaan kembali keterampilan dan meningkatkan kemampuan belajar dari sedikit data, terutama di lingkungan dengan imbalan tertunda atau tujuan kecil berurutan. Pendekatan representatif adalah HiLMDP, yang secara bersamaan mempelajari primitif motorik tersembunyi dan strategi utama atas primitif tersebut. Metode ini menunjukkan kemampuan berlaku pada keadaan lain yang kuat di seluruh tugas gerak berpindah dengan membagi strategi menjadi beberapa bagian.

Demikian pula, CompILE belajar untuk membagi contoh tindakan dari ahli menjadi segmen perilaku dengan panjang variabel secara tidak terawasi, memungkinkan sistem AI untuk meniru lintasan rumit dengan menyusun sub-keterampilan yang dapat diinterpretasikan. Metode-metode ini memiliki prinsip yang sama: dengan memisahkan apa yang harus dilakukan dari bagaimana melakukannya, sistem AI lebih siap untuk berlaku pada keadaan lain di berbagai tugas, mentransfer keterampilan, dan berpikir atas struktur berdasarkan waktu. Dengan demikian, IL bertingkat menyediakan kerangka kerja yang ampuh untuk membangun model perilaku yang mudah dikembangkan dan terstruktur, terutama di bidang robotika dan yang berwujud.

Perluasan dan Kemajuan dalam IL

Kemajuan terbaru dalam IL telah memperluas cakupannya di luar pengaturan biasa tugas tunggal yang sepenuhnya diawasi. Salah satu perluasan utama adalah peniruan dari observasi (IfO), di mana sistem AI belajar hanya dari lintasan keadaan ahli tanpa akses ke tindakan yang sesuai. Pengaturan ini sangat berguna dalam penggunaan di dunia nyata seperti contoh tindakan video atau catatan kendali jarak jauh manusia. Metode seperti *Time-Contrastive Networks* (TCN) dan cara terjemahan konteks mengatasi tantangan ini dengan mempelajari pengganti imbalan atau ruang bentuk angka ringkas yang selaras dengan perilaku ahli. Arah lain melibatkan pembelajaran dari contoh tindakan yang tidak sempurna atau kurang baik, yang umum terjadi dalam keadaan praktis. Untuk menangani data yang bising atau tidak tetap, metode terbaru memperkenalkan cara kerja penyaringan, rumus kesalahan berbobot kepercayaan, atau rumusan saling bersaing yang kuat. Misalnya, DAC memodifikasi kerangka kerja GAIL untuk mengakomodasi contoh tindakan berkualitas campuran dengan membentuk sinyal imbalan sesuai kebutuhan. Pendekatan ini meningkatkan keandalan peniruan di bawah pengawasan terbatas atau umpan balik yang tidak pasti.

Belajar dengan meniru interaktif dan langsung melalui jaringan lebih lanjut memperluas penggunaan IL ke sistem AI berwujud yang beroperasi di lingkungan langsung. Langkah kerja seperti DAgger dan penerusnya memungkinkan sistem AI untuk meminta masukan ahli selama pelaksanaan, secara bertahap meningkatkan daya tahan strategi melalui pengumpulan data berulang. IL juga telah dikombinasikan dengan cara lain, seperti RL dan belajar cara belajar, untuk meningkatkan kemampuan berlaku pada keadaan lain dan kemampuan menyesuaikan diri. Misalnya, metode seperti LfD+RL melatih strategi terlebih dahulu dengan contoh tindakan sebelum memperbaikinya melalui sinyal penguatan, memanfaatkan kekuatan kedua cara. Strategi hibrida ini sangat efektif dalam tugas dengan imbalan yang jarang atau pengaturan kritis keselamatan di mana eksplorasi dibatasi.

Batasan IL

Seiring IL terus berkembang, penerapannya dalam sistem kecerdasan berwujud memperkenalkan tantangan dan peluang baru. Metode IL tradisional sering mengasumsikan lingkungan tetap, pengamatan penuh, dan contoh tindakan berkualitas tinggi yang melimpah. Sebaliknya, pengaturan berwujud di dunia nyata bersifat berubah-ubah, sebagian dapat diamati, dan secara bawaan dengan banyak jenis data. Kerumitan ini telah memotivasi beberapa arah riset yang muncul yang mendorong batasan IL. Salah satu batasan utama menyangkut peniruan di bawah pengamatan tidak lengkap. Sistem AI berwujud sering beroperasi dengan masukan dari sensor yang terhalang, umpan balik yang tertunda, atau informasi keadaan yang tidak lengkap. Mengatasi hal ini membutuhkan penggabungan susunan sistem berbasis memori, model strategi berulang, dan cara kerja penarikan kesimpulan tersembunyi. Misalnya, kerangka kerja peniruan yang sadar POMDP menggabungkan penarikan kesimpulan Bayesian berbasis variasi dengan strategi berulang, memungkinkan sistem AI untuk mereplikasi perilaku ahli bahkan di bawah ketidakpastian pengamatan yang besar.

Arah aktif lainnya adalah Belajar dengan meniru dengan banyak jenis data, di mana contoh tindakan diberikan melalui beragam saluran dari sensor seperti penglihatan, bahasa, audio, dan rasa posisi tubuh. Hal ini sangat relevan dengan interaksi manusia-robot, di mana instruksi dapat disampaikan melalui bahasa alami atau isyarat. Karya terbaru telah berfokus pada penyesuaian agar selaras antarjenis data, penggabungan berdasarkan perhatian, dan strategi penyesuaian spesifik jenis data, yang semuanya berkontribusi pada pembelajaran yang lebih kuat dan luwes dalam kondisi dunia nyata.

Terakhir, kemajuan terkini berupaya meningkatkan kemampuan penyesuaian dan kemampuan berlaku pada keadaan lain IL di luar cara tetap dan sepenuhnya terawasi. Peniruan lintas bentuk fisik mengatasi tantangan mentransfer contoh tindakan di berbagai bentuk tubuh yang berbeda, yang membutuhkan pembelajaran gambaran perilaku yang tidak bergantung pada bentuk fisik. Belajar dengan meniru interaktif memungkinkan sistem AI untuk menyempurnakan strategi mereka melalui umpan balik ahli yang terbatas, daripada hanya mengandalkan lintasan yang telah dikumpulkan sebelumnya. Secara paralel, pendekatan yang sadar akan hubungan sebab-akibat bertujuan untuk mengungkap struktur sebab-akibat yang mendasari tindakan ahli, memudahkan transfer strategi ke pengaturan baru dengan menangkap maksud dan efek daripada perilaku dangkal. Secara bersama-sama, arah-arah ini mewakili langkah-langkah penting menuju sistem peniruan yang lebih mandiri, dapat digeneralisasi, dan sadar konteks untuk sistem AI berwujud.

Model AI Bahasa Besar dengan Banyak Jenis Data

Model AI Bahasa Besar dengan Banyak Jenis Data (MLLM) merupakan kemajuan besar penting dalam kecerdasan buatan masa kini. Dengan menggabungkan penglihatan, bahasa, audio, dan, dalam beberapa kasus, sinyal tindakan dalam cara kerja yang menyatu, model-model ini memperluas kemampuan model bahasa berdasarkan teks biasa untuk mendukung pengamatan, proses berpikir, dan interaksi di berbagai jenis data. Dalam sistem berwujud dan interaktif, MLLM berfungsi sebagai bagian sentral yang menjembatani pemahaman pemahaman makna tingkat umum dengan masukan gabungan sensor dan gerakan tingkat rendah. Pengembangannya telah memungkinkan interaksi manusia-mesin yang lebih lancar dan alami serta mendukung kelas model penglihatan-bahasa-tindakan (VLA) yang semakin banyak diterapkan dalam robotika dan lingkungan dunia nyata.

Taksonomi MLLM

MLLM memperluas kemampuan inti LLM dengan menggabungkan pengamatan antarjenis data seperti penglihatan, audio, dan tindakan. Secara susunan sistem, MLLM dibangun di atas kerangka kerja Transformer, ditambah dengan pengubah masukan untuk tiap jenis data, bagian penggabung antarjenis data, atau gambaran data tersembunyi terpadu. Dalam pelatihan, model-model ini mengadopsi strategi mulai dari pembelajaran dengan membandingkan hingga penyesuaian berdasarkan petunjuk dan penyelarasan melalui data buatan. Bidang ini telah mengalami konvergensi menuju susunan sistem yang menyatu yang mudah dikembangkan yang mendukung kemampuan berlaku pada keadaan lain tanpa contoh khusus dan penggunaan interaktif dunia nyata. Di bawah ini, kami meninjau tiga cara representatif yang telah membentuk pengembangan MLLM.

Susunan sistem berbasis Pembelajaran Dengan membandingkan. Susunan sistem berbasis pembelajaran dengan membandingkan adalah garis dasar untuk mencapai penyesuaian agar selaras jenis data tanpa memerlukan pengawasan jelas. CLIP adalah model representatif dalam kategori ini. Ia menggunakan struktur dua pengubah masukan yang secara terpisah mengkodekan gambar dan teks ke dalam ruang bentuk angka ringkas bersama dan melatihnya untuk mencocokkan di seluruh sampel berpasangan menggunakan nilai kesalahan perbandingan.

Desain ini memungkinkan pelatihan yang mudah dikembangkan pada kumpulan data bising skala web dan memungkinkan tugas klasifikasi dan pengambilan tanpa contoh khusus tanpa penyesuaian lanjutan khusus tugas. Ekstensi selanjutnya termasuk FILIP, yang menyempurnakan penyesuaian agar selaras tingkat potongan data antara potongan gambar dan bentuk angka ringkas teks, dan ALIGN, yang meningkatkan skala pelatihan gaya CLIP hingga miliaran pasangan gambar-teks bising, meningkatkan kemampuan berlaku pada keadaan lain di berbagai bidang gambar. Pendekatan kontras ini menekankan penyesuaian agar selaras makna pada tingkat gambaran dan berfungsi sebagai dasar bagi banyak sistem pengambilan dan pengaitan. Susunan sistem Transformer Dengan banyak jenis data Terpadu. Susunan sistem transformer dengan banyak jenis data terpadu bertujuan untuk menggabungkan beberapa jenis data masukan dalam satu tulang punggung Transformer, alih-alih merancang pengubah masukan untuk tiap jenis data. Contoh representatif adalah Perceiver IO, yang memperkenalkan bagian sempit tersembunyi dan cara memilih bagian penting silang untuk memproses masukan memiliki banyak unsur secara luwes di berbagai bidang. Desain ini memisahkan ukuran masukan dari kerumitan model dan memungkinkan proses berpikir dengan banyak jenis data yang mudah dikembangkan.

Berdasarkan ide ini, model seperti PaLI dan Flamingo lebih lanjut menunjukkan potensi tulang punggung terpadu. PaLI menggunakan Transformer pengubah masukan dan pembuat keluaran bersama untuk tugas bahasa-penglihatan dalam banyak bahasa, sementara Flamingo menyisipkan perhatian antarjenis data ke dalam LLM beku, memungkinkan penyesuaian dengan sedikit contoh untuk tugas bahasa-penglihatan. Susunan sistem ini telah terbukti efektif di berbagai tugas seperti pembuatan keterangan gambar, menjawab pertanyaan gambar, dan mengikuti instruksi, menekankan keluwesan dan efisiensi strategi pembuatan model terpadu.

Susunan sistem Penyesuaian Instruksi. Susunan sistem penyesuaian instruksi menyediakan strategi efektif lain untuk melengkapi LLM dengan kemampuan dengan banyak jenis data. LLaVA mewakili alur kerja representatif dalam kategori ini. Ia menghubungkan pengubah gambar beku dengan LLM yang telah dilatih sebelumnya melalui bagian penghubung yang dapat dipelajari, dan kemudian menyempurnakan model gabungan menggunakan pasangan instruksi-respons dengan banyak jenis data buatan yang dihasilkan oleh GPT-4. Pendekatan ini memungkinkan penyesuaian agar selaras penglihatan-bahasa dengan data spesifik tugas minimal dan mencapai kinerja yang kuat dalam menjawab pertanyaan gambar, pembuatan keterangan, dan dialog. Karya selanjutnya, seperti LLaVA-1.5 dan MiniGPT-4, lebih lanjut memperbaiki penghematan data, penyederhanaan susunan

sistem, dan kualitas instruksi. Dengan demikian, MLLM yang disetel instruksi memungkinkan penyesuaian yang mudah dikembangkan dari model yang berpusat pada bahasa menjadi sistem AI dengan banyak jenis data dengan kemampuan pengamatan dan proses berpikir yang mendasar.

Perluasan dan Kemajuan dalam MLLM

Perkembangan terkini telah menunjukkan penggabungan praktis Model AI Bahasa Besar dengan Banyak Jenis Data ke dalam sistem dunia nyata. Contoh yang menonjol adalah GPT-4V (GPT-4 dengan Penglihatan), yang dirilis oleh OpenAI pada akhir tahun 2023. GPT-4V memperluas susunan sistem GPT-4 dengan menerima masukan gambar bersamaan dengan perintah tekstual, memungkinkan berbagai aplikasi, termasuk pemahaman dokumen, penafsiran gambar medis, dan proses berpikir berdasarkan diagram. Kemampuannya lebih lanjut menyatu dalam GPT-4o, yang mendukung interaksi dengan banyak jenis data langsung di seluruh saluran teks, gambar, dan audio. GPT-4o menggabungkan sintesis dan pengenalan suara, dasar berdasarkan penglihatan, dan proses berpikir dengan banyak jenis data ke dalam satu lingkaran penarikan kesimpulan, menandai pergeseran menuju sistem AI yang terpadu, responsif, dan berwujud.

Keluarga model Gemini dari Google DeepMind mendorong batas-batas proses berpikir dengan banyak jenis data. Model Gemini dilatih pada data penglihatan-bahasa-audio skala besar dan selaras, dan menunjukkan kinerja yang kuat pada tugas dalam banyak bahasa dan konteks panjang, termasuk dialog gambar, penafsiran kode, dan pemahaman video. Melengkapi hal ini, seri Gemma berbobot terbuka memperluas kerangka kerja PaLI untuk mendukung proses berpikir penglihatan-bahasa tujuan umum dengan batas informasi panjang yang dapat dibaca sekaligus dan strategi penyempurnaan yang efisien. Dalam bidang AI berwujud, model Penglihatan-Bahasa-Tindakan (VLA) mencontohkan bentuk fisik MLLM. Model RT-X menyempurnakan model dasar penglihatan-bahasa pada lintasan robot berpasangan, memungkinkan kemampuan berlaku pada keadaan lain tanpa contoh khusus yang dikondisikan instruksi untuk tugas penggerakan benda di berbagai platform. Pendekatan ini telah diperluas dalam sistem seperti Helix dan Gemini Robotics, yang menggunakan desain terbagi menjadi beberapa bagian yang memisahkan perencanaan bahasa-penglihatan tingkat tinggi dari kendali gabungan penglihatan dan gerakan tingkat rendah, menghasilkan strategi robot yang mudah dikembangkan dan dapat ditransfer dalam keadaan dunia nyata.

Batasan MLLM

Seiring MLLM terus berkembang, beberapa batasan telah muncul di persimpangan kecerdasan dengan banyak jenis data, interaksi dunia nyata, dan proses berpikir berwujud. Perkembangan ini mencerminkan upaya yang lebih luas untuk memperluas MLLM di luar patokan uji tetap dan masukan yang dikurasi, mendorong ke arah sistem AI dunia terbuka interaktif yang mampu melakukan pengamatan, proses berpikir, dan tindakan di lingkungan berubah-ubah. Salah satu arah utama adalah penggabungan MLLM dengan sistem berwujud, memungkinkan proses berpikir pengamatan-bahasa-tindakan dalam platform robot. Model seperti RT-2 dan ChatVLA-2 menunjukkan bahwa penyempurnaan model bahasa-penglihatan yang telah dilatih sebelumnya pada data lintasan robot memungkinkan sistem AI untuk

menerapkan pada keadaan lain pengetahuan dari kumpulan data skala web ke tugas penggerakan benda fisik. Sistem ini menjembatani proses berpikir berdasarkan makna tingkat tinggi dan kendali motorik tingkat rendah, memudahkan kemampuan berlaku pada keadaan lain tanpa contoh khusus di dunia nyata. Kemajuan terbaru seperti Gemini Robotics dan Figure's Helix lebih lanjut mengadopsi cara sistem ganda, di mana MLLM tingkat tinggi memandu pengambilan keputusan sementara pengontrol tingkat rendah menangani ketepatan gabungan penglihatan dan gerakan.

Bidang kedua berpusat pada proses berpikir interaktif beberapa putaran antarjenis data. MLLM tradisional dilatih pada data tetap berpasangan, tetapi penggunaan di dunia nyata membutuhkan dialog beberapa langkah yang didasarkan pada pengamatan dan memori. Model seperti GPT-4o menggabungkan interaksi audio dan video langsung, memungkinkan percakapan dengan banyak jenis data sinkron, sementara riset dalam mengikuti instruksi (misalnya, LLaVA-1.5, PaliGemma) menekankan kaitan dengan urutan waktu dan proses berpikir tentang ruang dalam pengaturan interaktif. Kemampuan tersebut sangat penting untuk aplikasi seperti robotika bantu, sistem bimbingan, dan sistem AI mandiri yang beroperasi di ruang fisik yang rumit. Bidang aktif ketiga berfokus pada penyesuaian berkelanjutan dalam pengaturan dengan banyak jenis data. Saat sistem AI berinteraksi dengan dunia, mereka menemukan bidang, kosakata, atau distribusi gambar baru. Oleh karena itu, MLLM harus memperoleh kemampuan untuk menyesuaikan diri tanpa melupakan secara katastropik. Metode seperti finetuning berdasarkan memori, penyesuaian peringkat rendah (LoRA), dan penyetelan adaptor terbagi menjadi beberapa bagian menawarkan solusi yang menjanjikan, tetapi pembelajaran dengan banyak jenis data berkelanjutan tetap menjadi tantangan terbuka. Selain itu, upaya untuk menyelaraskan MLLM dengan prioritas sebab-akibat dan berwujud menekankan kebutuhan akan model yang tidak hanya menafsirkan tetapi juga bertindak secara bermakna.

Terakhir, penafsiran dan efisiensi muncul sebagai batas dasar. Ukuran dan kerumitan MLLM yang semakin besar menimbulkan kekhawatiran tentang sifat yang sulit dipahami dan konsumsi energi. Pendekatan seperti gabungan beberapa ahli yang dipilih seperlunya, pemberian petunjuk terstruktur, dan pengurangan potongan data bertujuan untuk meningkatkan efisiensi, sementara analisis perhatian berdasarkan tingkat kepentingan dan penyelidikan sebab-akibat sedang dieksplorasi untuk mengungkap bagaimana MLLM menggabungkan informasi dengan banyak jenis data. Seiring dengan perluasan penggunaan ke bidang yang kritis terhadap keselamatan, kekhawatiran ini menjadi semakin penting.

Model Dunia

Dalam pengembangan kecerdasan mandiri berwujud, kemampuan untuk membuat model, menirukan, dan mengantisipasi perubahan lingkungan secara internal merupakan kemampuan mendasar. Konsep model dunia tidak mengacu pada gambaran lingkungan yang tetap, tetapi pada cara kerja internal yang berubah-ubah dan terstruktur yang mendukung pengamatan, penarikan kesimpulan, dan pengambilan keputusan. Melalui model-model tersebut, sistem AI mampu belajar dari pengalaman yang dibayangkan, mengurangi ketergantungan pada interaksi dunia nyata, dan mengejar tujuan jangka panjang melalui

pertimbangan internal. Model dunia yang terbentuk dengan baik memungkinkan sistem AI untuk memperkirakan konsekuensi tindakannya, mempertahankan kesinambungan dari waktu ke waktu dalam proses berpikir, dan menghasilkan strategi perilaku yang efisien bahkan dalam kondisi ketidakpastian.

Asal usul filosofis model dunia dapat ditelusuri ke pernyataan Immanuel Kant bahwa pemahaman muncul dari gambaran di dalam sistem daripada sekadar penerimaan masukan dari sensor. Gagasan ini kemudian diperluas oleh para ahli teori terkait proses berpikir, termasuk Jay Forrester, yang mengusulkan bahwa individu bergantung pada tiruan di dalam sistem, atau model mental, untuk memandu pengambilan keputusan. Ilmu saraf masa kini mendukung perspektif ini, mengungkapkan bahwa otak membangun dan memperbarui gambaran umum ruang dan waktu untuk memperkirakan peristiwa masa depan dan mengatur perilaku. Dalam kecerdasan buatan, dasar berdasarkan teori ini telah memotivasi desain model yang mempelajari pendekatan berdasarkan peluang terhadap perubahan lingkungan. Model-model ini biasanya terdiri atas aturan perubahan yang memperkirakan distribusi $p(s_{t+1} \mid s_t, a_t)$ dan, dalam banyak kasus, juga mencakup aturan imbalan $r(s_t, a_t)$. Model internal tersebut memberi sistem AI kemampuan untuk melakukan tiruan rangkaian langkah di dalam sistem, memungkinkan perencanaan yang efisien, kendali yang terinformasi, dan pembelajaran strategi yang hemat data.

Sebagian besar penerapan model dunia mengadopsi struktur terbagi menjadi beberapa bagian yang terdiri dari tiga bagian inti: pengubah data pengamatan, bagian penyimpan urutan waktu, dan pengontrol strategi. Pengubah masukan mengubah masukan dari sensor memiliki banyak unsur menjadi gambaran data tersembunyi yang ringkas, mempertahankan informasi penting sambil membuang variasi yang tidak relevan. Unit memori menangkap ketergantungan berdasarkan waktu dengan memperbarui kondisi tersembunyi berdasarkan pengamatan dan tindakan masa lalu. Kondisi tersembunyi ini merangkum riwayat sistem AI dan mendukung perkiraan tentang keadaan di masa depan. Pengontrol kemudian memilih tindakan berdasarkan keadaan memori dan pengamatan saat ini, memandu interaksi fisik dengan lingkungan atau lintasan yang dibayangkan yang digunakan untuk perencanaan. Bagian-bagian ini diperbaiki secara bersamaan menggunakan nilai kesalahan berdasarkan perkiraan untuk observasi dan imbalan, seringkali dilengkapi dengan tujuan berbasis variasi untuk mengatur ruang data tersembunyi. Bersama-sama, mereka membentuk susunan sistem tergabung yang memungkinkan sistem AI mandiri untuk menirukan hasil, berpikir atas urutan berdasarkan waktu, dan menyesuaikan diri dengan situasi baru tanpa pengumpulan data dunia nyata yang ekstensif.

Model kondisi berulang (RSSM)

Model kondisi berulang menyediakan susunan sistem yang banyak diadopsi untuk membuat model perubahan tersembunyi dalam lingkungan Belajar melalui imbalan di mana pengamatan bersifat parsial dan ketergantungan berurutan bersifat jangka panjang. RSSM menangani dua persyaratan utama dalam pengaturan tersebut: kebutuhan untuk mempertahankan memori urutan waktu yang selaras dan kebutuhan untuk membuat model ketidakpastian dalam pengamatan dan perpindahan. Untuk memenuhi persyaratan ini, RSSM

memperkenalkan struktur tersembunyi hibrida yang terdiri dari kondisi tersembunyi yang pasti dan nilai tersembunyi yang acak. Desain ini memungkinkan model untuk memisahkan memori terstruktur dari perkiraan ketidakpastian, sehingga meningkatkan ketepatan perkiraan dan kemampuan berlaku pada keadaan lain.

Pada setiap langkah waktu, model menerima kondisi tersembunyi sebelumnya, keadaan acak sebelumnya, dan tindakan sebelumnya. Ketiganya digabungkan untuk menghasilkan kondisi tersembunyi yang diperbarui melalui fungsi berulang, yang diformalkan sebagai $h_t = f(h_{t-1}, z_{t-1}, a_{t-1})$. Berdasarkan kondisi tersembunyi ini dan pengamatan saat ini, model menyimpulkan distribusi perkiraan setelah melihat data atas nilai tersembunyi baru $z_t \sim q(z_t \mid h_t, o_t)$, sambil juga mempertahankan distribusi perkiraan awal yang dipelajari $p(z_t \mid h_t)$. Hal ini memungkinkan model menghasilkan perkiraan selama imajinasi ketika tidak ada pengamatan baru yang tersedia. Pasangan (h_t, z_t) digunakan untuk merekonstruksi pengamatan, memperkirakan imbalan, dan memperkirakan apakah episode saat ini telah berakhir. Model ini dilatih dengan membuat sekecil mungkin rumus kesalahan berbasis variasi yang menyeimbangkan ketepatan pembentukan ulang dan pembatasan, yang biasanya dinyatakan sebagai jumlah dari suku nilai kecocokan data dan perbedaan Kullback-Leibler antara perkiraan setelah melihat data dan perkiraan awal.

Kondisi tersembunyi yang pasti bertindak sebagai memori jangka panjang, memungkinkan model untuk mengkodekan urutan dengan konsistensi berdasarkan waktu. Nilai tersembunyi yang acak, sebaliknya, memungkinkan model untuk menggambarkan ketidakjelasan dalam pengamatan dan perpindahan, yang penting ketika lingkungan bising atau sebagian dapat diamati. Pembagian fungsi ini meningkatkan stabilitas selama pelatihan dan menghasilkan perencanaan yang lebih kuat. Setelah pelatihan, RSSM mendukung perencanaan berdasarkan imajinasi dengan mengambil sampel dari perkiraan awal dan memperbarui kondisi tersembunyi di ruang data tersembunyi tanpa memerlukan pengamatan baru. Tiruan rangkaian langkah ini kemudian dapat digunakan untuk memperbarui strategi dan rumus nilai melalui Belajar melalui imbalan, secara substansial meningkatkan kemampuan belajar dari sedikit data dan mengurangi beban interaksi lingkungan.

Berbagai perluasan RSSM telah diusulkan untuk meningkatkan penerapannya. Versi bertingkat memperkenalkan beberapa lapisan nilai tersembunyi untuk menangkap informasi pada tingkat abstraksi yang berbeda, meningkatkan kinerja dalam keadaan banyak tugas atau pembelajaran transfer. Varian lain menggabungkan perubahan waktu berkelanjutan menggunakan persamaan perubahan biasa neural untuk menangani data yang diambil sampelnya secara tidak teratur. Terlepas dari rumusan spesifiknya, wawasan utama RSSM tetap sama: dengan menggabungkan memori terstruktur dan penarikan kesimpulan berdasarkan peluang dalam cara kerja yang menyatu, ia memungkinkan pembuatan model pengalaman berurutan yang luwes dan akurat pada sistem AI mandiri.

Susunan sistem perkiraan berbasis bentuk angka ringkas

Susunan sistem perkiraan berbasis bentuk angka ringkas Bersama memperkenalkan pendekatan yang secara dasar berbeda untuk pembuatan model dunia dengan menggeser tujuan pembuatan model dari pembentukan ulang masukan dari sensor mentah ke perkiraan

dalam ruang bentuk angka ringkas yang umum. JEPA menghindari kerumitan pembentukan ulang tingkat piksel atau tingkat sinyal dan sebagai gantinya berfokus pada penyesuaian agar selaras gambaran data tersembunyi dari keadaan saat ini dan masa depan. Penekanan pada penyesuaian agar selaras gambaran ini mencerminkan pengamatan bahwa informasi makna dan struktural tingkat tinggi seringkali lebih relevan untuk pengambilan keputusan daripada detail dari sensor tingkat rendah. JEPA terdiri dari enam bagian sistem yang berbeda secara fungsional.

Pengatur menentukan target perkiraan dan horizon perencanaan. Pengamat mengkodekan pengamatan mentah ke dalam bentuk angka ringkas tersembunyi yang berfungsi sebagai gambaran keadaan internal. Model dunia memperkirakan bentuk angka ringkas masa depan berdasarkan bentuk angka ringkas saat ini dan memori internal. Penilai biaya memberi skor pada perkiraan ini berdasarkan tujuan tugas atau kriteria yang dipelajari. Perencana tindakan memilih sinyal kendali yang membuat sekecil mungkin biaya yang diharapkan di bawah perubahan yang diperkirakan. Memori jangka pendek mempertahankan riwayat yang relevan untuk mendukung penarikan kesimpulan berurutan. Bersama-sama, Bagian-bagian ini membentuk sistem perkiraan dari awal sampai akhir yang mampu melakukan simulasi ke depan, penilaian biaya, dan perencanaan strategi di ruang data tersembunyi.

Prinsip inti yang mendasari JEPA adalah penyesuaian bentuk angka ringkas untuk membuat perkiraan. Dengan bentuk angka ringkas saat ini, model mempelajari fungsi untuk memperkirakan bentuk angka ringkas di masa mendatang, dan perkiraan ini dibandingkan dengan bentuk angka ringkas masa depan yang sebenarnya yang diperoleh dari pengamatan selanjutnya. Rumus kesalahan yang digunakan untuk mengukur perbedaan ini dapat mengambil berbagai bentuk, termasuk kesalahan kuadrat rata-rata, kesamaan kosinus, atau perbedaan kontras. Tujuan pelatihan ini mendorong model untuk mengekstrak fitur yang stabil secara berdasarkan waktu dan relevan secara sebab-akibat, seperti permanensi objek atau pola gerakan, sambil menghindari terlalu menyesuaikan diri dengan data latihan terhadap variasi dari sensor yang tidak relevan.

Dengan menghindari overhead perhitungan komputer dari generasi resolusi tinggi dan berfokus pada abstraksi yang relevan dengan tugas, JEPA meningkatkan efisiensi pelatihan dan kemampuan berlaku pada keadaan lain. Selain itu, rumusan untuk memperkirakan latennya memungkinkan penggabungan yang mulus dengan bagian sistem perencanaan dan kendali, yang dapat beroperasi langsung pada bentuk angka ringkas tanpa memerlukan pengubahan kode menjadi keluaran jelas. JEPA juga mendukung pelatihan terbagi menjadi beberapa bagian dan pemindahan hasil belajar, karena setiap bagian dapat dilatih secara independen dan digunakan kembali dalam tugas-tugas baru. Susunan sistem ini selaras dengan teori pemrosesan untuk memperkirakan dalam ilmu saraf, yang menyatakan bahwa sistem biologis terus-menerus mengantisipasi keadaan masa depan dalam ruang gambaran di dalam sistem dan menyesuaikan perilaku sesuai dengan itu. JEPA menyediakan struktur perhitungan komputer formal untuk cara kerja tersebut pada sistem AI buatan.

Model dunia pembuat data

Model dunia pembuat data dirancang untuk memperkirakan pengamatan di masa depan yang dikondisikan pada keadaan dan tindakan saat ini, dengan tujuan menghasilkan hasil dari sensor yang realistis dan selaras. Model-model ini belajar untuk mendekati distribusi pengamatan di masa depan atau pengkodean latennya, memungkinkan simulasi perubahan lingkungan dengan cara yang didasarkan pada pengamatan. Kemampuan seperti itu sangat berharga dalam tugas-tugas di mana detail gambar, penggabungan dengan banyak jenis data, atau penafsiran manusia sangat penting untuk pengambilan keputusan yang efektif. Model dunia pembuat data standar terdiri dari tiga bagian: pengubah masukan yang memetakan pengamatan ke nilai tersembunyi, model perpindahan yang memperkirakan kondisi tersembunyi berikutnya berdasarkan kondisi tersembunyi dan tindakan saat ini, dan pembuat keluaran yang merekonstruksi pengamatan yang diperkirakan. Model dilatih menggunakan nilai kesalahan pembentukan ulang yang dikombinasikan dengan istilah pembatasan yang membatasi ruang data tersembunyi, seringkali menggunakan cara penarikan kesimpulan berbasis variasi. Hasilnya adalah model yang mampu menghasilkan urutan pengamatan di masa depan yang selaras secara struktural dan makna.

Karya terbaru telah memperluas model yang menghasilkan data ke bidang yang lebih rumit. Model yang menghasilkan data dengan banyak jenis data seperti MUVO menggabungkan masukan gambar dengan informasi ruang untuk memperkirakan peta keterisian ruang tiga dimensi, mendukung navigasi yang aman dan terarah. Model yang dikondisikan bahasa menggabungkan informasi tekstual ke dalam proses perpindahan tersembunyi, memungkinkan sistem AI untuk menirukan hasil sesuai dengan instruksi verbal atau konteks dialog. Dalam model seperti Language World Model, lintasan tersembunyi digunakan untuk menghasilkan deskripsi bahasa alami tentang masa depan yang dibayangkan, memungkinkan sistem AI untuk merefleksikan rencana internal dan mengkomunikasikan perkiraan kepada pengguna manusia. Model yang menghasilkan data memberikan beberapa keuntungan penting. Mereka menawarkan akses jelas ke hasil dari sensor yang diperkirakan, yang dapat digunakan untuk memandu kendali berdasarkan penglihatan, melakukan proses berpikir tentang kemungkinan yang berbeda, atau menghasilkan penjelasan yang selaras dengan manusia. Namun, mereka juga menghadirkan tantangan. Generasi resolusi tinggi membutuhkan perhitungan komputer yang intensif dan seringkali tidak stabil, dan sintesis akurat dari pengamatan dengan banyak jenis data yang rumit tetap menjadi masalah riset yang terbuka. Untuk mengatasi masalah ini, karya terbaru telah mengeksplorasi susunan sistem hibrida yang menggabungkan proses berpikir ruang data tersembunyi dengan pengubahan kode menjadi keluaran yang menghasilkan data, memungkinkan sistem AI untuk merencanakan secara efisien sambil mempertahankan kemampuan untuk memvisualisasikan atau memverbalisasikan proses internal mereka.

Batasan Model Dunia

Model dunia berfungsi sebagai dasar internal untuk pengamatan, imajinasi, dan pengambilan keputusan pada sistem AI mandiri yang berwujud. Model ini menyediakan cara kerja struktural yang memungkinkan sistem AI untuk memahami konsekuensi tindakan

mereka, belajar dari pengalaman simulasi, dan memperbaiki perilaku dengan mengurangi ketergantungan pada umpan balik eksternal. Di berbagai cara pembuatan model, dari struktur urutan waktu gabungan RSSM, hingga pembelajaran gambaran umum JEPA, dan kelengkapan pengamatan model pembuat data, muncul tujuan umum: membangun gambaran di dalam sistem tentang lingkungan yang mampu membuat perkiraan, berlaku pada keadaan lain, dan memiliki makna yang jelas. Seiring model-model ini terus berkembang, integrasinya ke dalam susunan sistem AI yang menyatu menjanjikan kemajuan dalam otonomi, keluwesan, dan penafsiran sistem cerdas. Dengan menggabungkan kekuatan di seluruh susunan sistem (misalnya, kesinambungan dari waktu ke waktu RSSM, penyederhanaan data tersembunyi JEPA, dan kemampuan yang luwes model yang menghasilkan data), sistem di masa depan mungkin dapat berpikir tentang gagasan umum, bertindak sesuai perubahan, dan berinteraksi secara jelas dengan manusia dan dunia fisik. Dalam konteks ini, model dunia bukan hanya alat perhitungan komputer, tetapi juga penghubung gagasan yang menghubungkan pengamatan, proses berpikir, dan tindakan dalam sistem AI cerdas.

1.3 KOMPONEN UTAMA AI BERWUJUD

Didorong oleh kemajuan dalam teknologi-teknologi utama ini, AI berwujud telah mengalami kemajuan pesat. Berikut ini, kami menyajikan tinjauan terstruktur tentang perkembangan dalam tiga bagian utama.

Pengamatan Aktif

Pengamatan Aktif mengacu pada proses peningkatan perolehan informasi dari data lingkungan melalui tindakan yang disengaja dan selektif. Pada intinya, untuk memahami dunia secara lebih efektif, sistem AI secara aktif bergerak dan menjelajahi lingkungannya. Dengan terus-menerus mengambil sampel lingkungan melalui sistem sensoriknya, sistem AI membangun pemahaman berdasarkan pengamatan yang pada dasarnya didasarkan pada interaksi dan tindakan yang bertujuan. Dalam kerangka kerja ini, penafsiran data dari sensor secara bawaan terkait dengan perilaku yang diperlukan untuk memperolehnya. Pengamatan dan tindakan beroperasi dalam koordinasi yang erat untuk mendukung pengambilan keputusan dapat menyesuaikan diri. Konsep ini telah dieksplorasi secara luas di ranah gambar, di mana sistem AI seperti hewan, manusia, robot, atau platform kamera terus mengubah posisi mereka untuk meningkatkan pandangan terhadap objek tertentu atau menggunakan gerakan untuk merasakan struktur lingkungan, seperti menavigasi di sekitar rintangan atau mengenali batas objek.

Dari sudut pandang yang lebih formal, Pengamatan Aktif terletak pada persimpangan pembuatan model berdasarkan pengamatan dan strategi kendali. Pembuatan model mengacu pada gambaran sensor, bagian sistem pemrosesan, dan interaksi mereka. Model-model ini biasanya dikategorikan sebagai lokal atau global, tergantung pada jangkauan ruang dan temporalnya. Model lokal mengkarakterisasi nilai pengatur dan proses seperti distorsi lensa, panjang fokus, resolusi ruang, dan penyaringan band-pass. Sebaliknya, model global menggambarkan kinerja sistem secara keseluruhan dan memperkirakan interaksi antar bagian. Strategi kendali dalam Pengamatan Aktif melibatkan pengenalan urutan tindakan yang

membuat sekecil mungkin rumus kesalahan sambil membuat sebesar mungkin perolehan informasi. Studi sebelumnya telah menunjukkan kelayakan rumusan tersebut, termasuk tugas-tugas seperti pembagian gambar 2D dan parameterisasi bentuk 3D. Pendekatan ini telah memungkinkan sistem AI untuk memperkirakan kedalaman objek menggunakan isyarat fokus dan konvergensi dua kamera.

Definisi Pengamatan Aktif yang terkait erat namun lebih sempit menganggap pengamatan dan tindakan di dalam otak pada dasarnya menyatu. Menurut pandangan ini, ketika seseorang mengamati suatu tindakan, otak secara internal mensimulasikannya sebagai respons motorik potensial, menafsirkan perilaku yang diamati dalam konteks yang dapat dilaksanakan. Perspektif ini mendukung gagasan bahwa manusia dan hewan dapat mempelajari perilaku yang tepat dengan mengamati orang lain, menekankan penggabungan mendalam antara pengamatan dan tindakan dalam sistem biologis. Secara keseluruhan, Pengamatan Aktif mengacu pada sistem AI yang secara selektif memperoleh informasi dari pengamatan lingkungan. Metode Pengamatan Aktif yang ada dapat dibagi secara kasar menjadi tiga kategori: SLAM gambar, Memahami Ruang 3D, dan Penjelajahan Lingkungan Secara Aktif. Untuk menawarkan perspektif yang efektif tentang pendekatan Pengamatan Aktif, seperti yang dirangkum dalam Tabel 1.1, kami menganalisis metode representatif di sepanjang tiga dimensi praktis: jenis sensor, jenis fitur, dan keadaan yang berlaku.

Gambar SLAM

Simultaneous Localization and Mapping (SLAM) adalah teknologi penting yang memungkinkan sistem AI untuk melokalisasi diri mereka sendiri dan membangun peta lingkungan di lingkungan yang tidak dikenal. Masalah SLAM berdasarkan peluang bermula pada Konferensi Robotika dan Otomasi IEEE tahun 1986, di mana para ahli mengusulkan penggunaan metode teori perkiraan untuk tantangan pemetaan dan penentuan posisi. Awalnya dikembangkan dalam bidang robotika, SLAM bertujuan untuk memungkinkan sistem AI membangun peta lingkungan sekitarnya secara real-time hanya menggunakan data sensor, semuanya tanpa pengetahuan sebelumnya tentang lingkungan. Pada saat yang sama, sistem AI memperkirakan posisinya sendiri relatif terhadap peta yang berkembang.

Jalur Pengembangan dari SLAM ke VSLAM. Sistem SLAM dapat memanfaatkan berbagai sensor untuk mengumpulkan data dari lingkungan, termasuk sensor berdasarkan laser, sensor akustik, dan sensor gambar. Di antara ini, sensor gambar mencakup berbagai perangkat pencitraan, seperti kamera satu kamera, dua kamera, berdasarkan peristiwa, omnidirectional, dan RGB-D. Robot yang dilengkapi dengan sensor gambar mengandalkan data gambar yang ditangkap oleh kamera ini untuk memperkirakan posisi dan orientasinya relatif terhadap lingkungan sekitarnya. Proses melakukan SLAM menggunakan masukan gambar secara khusus disebut sebagai Gambar SLAM (VSLAM). Memanfaatkan data gambar dalam SLAM menawarkan beberapa keuntungan: persyaratan perangkat keras yang lebih rendah, deteksi dan pelacakan objek yang lebih intuitif, dan akses ke informasi gambar dan makna yang kaya. Gambar atau frame video yang ditangkap juga dapat digunakan kembali untuk tugas berdasarkan penglihatan lainnya seperti pembagian objek berdasarkan makna dan deteksi objek, karena menyimpan banyak data yang dapat diproses lebih lanjut. Fitur-fitur ini telah

menjadikan VSLAM sebagai topik penting dalam komunitas robotika dan telah mendorong riset dan pengembangan yang ekstensif oleh para ahli di bidang robotika dan penglihatan komputer selama beberapa dekade terakhir. Akibatnya, VSLAM dapat diterapkan pada berbagai keadaan yang membutuhkan pembentukan ulang lingkungan 3D, seperti robotika mandiri, realitas tertambah (AR), dan robotika layanan.

Tabel 1.1 Perbandingan tiga kategori metode Pengamatan Aktif. (PC: *Awan Titik*, G: Geometris, V: Volumetrik, S: Semantik, M: Multimodal, T: Topologis, SC: Pemahaman Adegan)

Metode	Tahun	Jenis sensor	Jenis fitur	Keadaan yang dapat diterapkan
SLAM Gambar				
CoSLAM	2012	RGB-D	G + V	SLAM Berubah-ubah
SLAM++	2013	RGB-D	S	Pemetaan tingkat objek
ORB-SLAM	2015	RGB-D + Dua kamera	G	SLAM Berubah-ubah
DS-SLAM	2018	RGB-D	G + S	SLAM Berubah-ubah
TwistSLAM	2022	RGB-D + Dua kamera	G + S	SLAM Berubah-ubah
GS-SLAM	2024	RGB-D	V	Pemetaan tingkat objek
Memahami Ruang 3D				
Gaudi	2022	RGB	V	SC Umum
Clip2Scene	2023	RGB + PC	M	SC yang dipandu bahasa
OpenScene	2023	RGB + PC	M	SC Umum
Lexicon3D	2024	RGB-D	S	SC yang dipandu bahasa
GraphDreamer	2024	RGB	T + S	SC Terstruktur
HUGS	2024	RGB-D	M	SC Umum
RegionPLC	2024	RGB + PC	M	SC yang dipandu bahasa
Penjelajahan Lingkungan Secara Aktif				
MAX	2019	RGB	S	Eksplorasi yang dipandu makna
Active Neural SLAM	2020	RGB-D	V	Eksplorasi berdasarkan bentuk dan ukuran ruang
APT	2021	RGB	S	Eksplorasi yang dipandu makna
Conan	2023	RGB	T	Eksplorasi berdasarkan bentuk dan ukuran ruang
DBMI^2-BPI	2023	RGB-D	V	Eksplorasi berdasarkan bentuk dan ukuran ruang
ActiveRIR	2024	RGB + Audio	M	Pengamatan Aktif antarjenis data

Bagian Dasar VSLAM. Mengikuti susunan sistem standar yang diperkenalkan oleh, sebagian besar metode SLAM terdiri dari dua thread pemrosesan paralel: pelacakan dan pemetaan. Salah satu cara dasar untuk mengkategorikan langkah kerja VSLAM didasarkan pada pendekatan dan strategi berbeda yang diadopsi dalam kedua thread ini. Berdasarkan jenis data yang digunakan, metode VSLAM biasanya dibagi menjadi dua kategori utama: metode langsung dan metode tidak langsung (berdasarkan fitur). Metode tidak langsung mengekstrak titik fitur (atau titik kunci) dari tekstur adegan dan melacaknya di seluruh frame dengan mencocokkan deskriptornya.

Meskipun metode ini efisien secara perhitungan komputer dalam tahap pengambilan dan pencocokan fitur, metode ini menunjukkan daya tahan dan ketepatan yang kuat di bawah variasi fotometrik dalam intensitas frame. Di sisi lain, metode langsung memperkirakan gerakan kamera langsung dari nilai intensitas piksel, membangun masalah perbaikan untuk membuat sekecil mungkin kesalahan fotometrik. Metode ini bergantung pada prinsip fotogrametri, menggunakan semua piksel keluaran dari kamera dan melacak korespondensinya di seluruh frame dengan membuat model batasan seperti konstansi kecerahan dan konsistensi warna. Karakteristik tersebut memungkinkan metode langsung untuk menangkap informasi yang lebih detail dari gambar, sehingga memungkinkan pembentukan ulang ruang 3D dengan ketepatan lebih tinggi.

Namun, meskipun metode langsung berkinerja lebih baik di lingkungan tanpa tekstur dan menghindari beban perhitungan komputer pengambilan ciri, metode ini sering menghadapi tantangan perbaikan skala besar. Kekuatan dan kelemahan masing-masing pendekatan telah mendorong para ahli untuk mengeksplorasi metode hibrida yang menggabungkan keduanya. Pendekatan ini biasanya menggabungkan tahap deteksi metode langsung dan tidak langsung, di mana satu fase menginisialisasi atau mengoreksi fase lainnya untuk mencapai kinerja keseluruhan yang lebih baik. Selain itu, karena sistem VSLAM tipikal terdiri dari front-end gambar odometry (VO) untuk memperkirakan gerakan lokal kamera dan back-end SLAM untuk memperbaiki peta yang dibangun, terdapat beragam cara penerapan Bagian-bagian ini di berbagai pendekatan.

VO memperkirakan pose robot berdasarkan konsistensi frame-ke-frame lokal dan meneruskan hasilnya ke back-end untuk perbaikan global. Dengan demikian, perbedaan utama antara VSLAM dan VO terletak pada apakah konsistensi global peta dan lintasan dipertimbangkan. Sistem VSLAM canggih seringkali mencakup dua bagian sistem tambahan: deteksi penutupan loop dan pembentukan ulang peta. Bagian-bagian ini mengenali lokasi yang sebelumnya dikunjungi, memungkinkan koreksi lintasan yang lebih akurat dan penyempurnaan peta berdasarkan pose kamera yang diperbarui. Sebagai teknologi dasar Pengamatan Aktif, SLAM gambar telah dipelajari secara ekstensif. Menurut Wang dan Kawan-kawan, metode yang ada termasuk dalam kelompok berdasarkan bentuk, ukuran, dan makna. Metode berdasarkan bentuk dan ukuran memanfaatkan isyarat ruang atau berdasarkan waktu, seperti aliran adegan padat, konsistensi triangulasi, dan struktur grafik, berkinerja baik dalam pengaturan tetap tetapi kesulitan dengan adegan berubah-ubah. Sebaliknya, metode

makna meningkatkan penentuan posisi dan pemetaan di lingkungan berubah-ubah dengan memanfaatkan informasi tingkat tinggi.

VSLAM Makna. VSLAM telah mengalami kemajuan pesat karena keunggulannya, seperti biaya sensor yang rendah, kemudahan penggabungan dengan jenis data penginderaan lainnya, dan kemampuan untuk menangkap informasi lingkungan yang kaya. Metode SLAM berdasarkan penglihatan tradisional telah mencapai hasil yang luar biasa; namun, kinerjanya dapat menurun di lingkungan yang rumit atau menantang. Pengembangan pembelajaran mendalam telah dengan jelas memajukan penglihatan komputer dan telah menarik perhatian yang semakin besar dalam penerapannya pada SLAM. Dalam konteks ini, penggabungan informasi makna, yang memberikan pemahaman tingkat tinggi tentang lingkungan, menjadi semakin penting. Dengan memanfaatkan isyarat makna, robot dapat lebih efektif menafsirkan lingkungan sekitarnya dan meningkatkan kekokohan dan kecerdasan sistem SLAM. Metode awal yang representatif meliputi SLAM++, yang menggabungkan makna tingkat objek, dan DS-SLAM, yang menerapkan pembelajaran mendalam untuk pemahaman adegan berubah-ubah. Model terbaru seperti TwistSLAM dan GS-SLAM lebih meningkatkan kekokohan dengan menggabungkan perbaikan berdasarkan bentuk dan ukuran dengan pembuatan model makna atau yang menghasilkan data.

Secara umum, metode VSLAM makna terdiri dari empat bagian utama berikut: Bagian sistem Pelacakan: Bagian sistem ini memperkirakan pose kamera dan membangun titik peta 3D menggunakan titik fitur 2D yang diekstrak dari frame video berurutan. Melalui perkiraan pose dan konstruksi titik 3D, bagian sistem ini membangun dasar untuk penentuan posisi dan pemetaan. Bagian sistem Pemetaan Lokal: Dengan memproses dua frame video berurutan, bagian sistem ini menghasilkan titik peta 3D baru. Titik-titik ini, bersama dengan bagian sistem penyesuaian bundel (BA), digunakan untuk menyempurnakan perkiraan pose kamera. Bagian sistem Penutupan Loop: Bagian sistem ini membandingkan keyframe berdasarkan fitur gambar yang diekstrak dan menilai kemiripannya untuk mendeteksi lokasi yang pernah dikunjungi sebelumnya. Setelah loop terdeteksi, bagian sistem ini menyesuaikan lintasan kamera dan memperbaiki peta yang dibangun untuk konsistensi global. Penghapusan Konteks Non-Kaku (NRCC): Tujuan utama NRCC adalah untuk menyaring objek yang berubah-ubah secara berdasarkan waktu dari frame video, sehingga mengurangi dampak negatifnya pada tahap penentuan posisi dan pemetaan. Hal ini terutama melibatkan proses masking atau pembagian untuk mengisolasi instance yang tidak stabil dalam adegan, seperti manusia. Dengan mengurangi jumlah titik fitur yang terlibat dalam proses perbaikan, NRCC menyederhanakan perhitungan komputer dan meningkatkan daya tahan sistem. Dengan menggabungkan informasi tingkat makna ke dalam VSLAM, metode ini dapat mengurangi ketidakpastian dalam perkiraan pose dan pembentukan ulang peta. Namun, tantangan utama yang sedang berlangsung terletak pada bagaimana memanfaatkan informasi makna yang diekstrak secara efektif tanpa menimbulkan beban perhitungan komputer yang besar.

Memahami Ruang 3D

Pemahaman adegan berfokus pada memungkinkan sistem AI untuk mempersepsikan, mensegmentasi, dan berpikir tentang lingkungan yang rumit dengan cara yang terstruktur dan

bermakna secara makna. Karya-karya terbaru telah memajukan bidang ini dengan menggabungkan model bahasa-penglihatan dan prioritas yang menghasilkan data. Upaya awal seperti Gaudi memperkenalkan model yang menghasilkan data untuk sintesis adegan sadar 3D. Clip2Scene dan OpenScene memanfaatkan bentuk angka ringkas bahasa-penglihatan untuk memudahkan pemahaman 3D yang efisien label dan kosakata terbuka. Pemahaman adegan terstruktur semakin ditingkatkan oleh Lexicon3D dan GraphDreamer, yang membuat model relasi tingkat objek dalam ruang 3D melalui gambaran terstruktur seperti grafik adegan atau leksikon makna. Sementara itu, cara pengaitan dengan banyak jenis data tingkat wilayah, yang dicontohkan oleh HUGS dan RegionPLC, menggabungkan petunjuk dan pengaitan ruang untuk mencapai pengamatan 3D yang terperinci dan terkondisi tujuan. Metode-metode ini memajukan pemahaman 3D holistik yang selaras dengan bahasa.

Deteksi Objek 3D. Riset terbaru tentang deteksi objek 3D secara luas dapat dikategorikan menjadi tiga pendekatan utama: metode berdasarkan proyeksi kamera-LiDAR, metode berdasarkan perhatian, dan metode berdasarkan Transformer antarjenis data. Karena sensor LiDAR dan kamera RGB beroperasi dalam sistem koordinat yang berbeda, penyesuaian kamera-LiDAR memerlukan penegakan konsistensi berdasarkan bentuk dan ukuran antara kumpulan titik 3D 3D dan gambar multi-pandangan 2D. Untuk menyelaraskan fitur titik 3D dengan fitur piksel 2D, metode berdasarkan proyeksi telah diusulkan. Pendekatan ini menggunakan intrinsik kamera (misalnya, panjang fokus, titik utama) dan ekstrinsik (misalnya, posisi, orientasi) untuk memproyeksikan data gambar 2D ke ruang 3D, memungkinkan korespondensi satu-ke-satu antara titik 3D dan fitur piksel yang diproyeksikan.

Kerangka kerja awal seperti PointPainting menghasilkan skor pembagian objek berdasarkan makna 2D dan memproyeksikannya ke titik 3D untuk meningkatkan gambaran kumpulan titik 3D untuk deteksi hanya LiDAR. Metode terbaru seperti EPNet dan EPNet++ memperkenalkan rumus kesalahan baru untuk meningkatkan konsistensi dalam klasifikasi, penentuan posisi, dan makna antarjenis data. PointAugmenting lebih lanjut meningkatkan fitur titik menggunakan isyarat 2D yang berasal dari CNN, yang lebih kuat daripada masker pembagian. MSMDFusion memperkenalkan titik virtual dengan mengangkat piksel gambar 2D dengan informasi makna ke ruang kotak kecil 3D melalui penyelesaian kedalaman. Titik virtual ini dipilih dan dikelompokkan sesuai kebutuhan untuk penggabungan dengan banyak jenis data yang ditingkatkan. Meskipun metode ini unggul dalam mendeteksi objek besar di dekat sensor LiDAR, mengenali objek kecil atau jauh tetap menjadi tantangan karena kepadatan titik yang jarang. Untuk mengatasi hal ini, metode seperti Frustum-PointPillars dan Faraway Frustum menggunakan masker instance 2D atau kotak pembatas untuk membangun frustum 3D yang melokalisasi objek yang jauh. MVP menambah kumpulan titik 3D dengan menghasilkan titik virtual padat di sekitar frustum latar depan untuk mengurangi perbedaan antara gambaran objek dekat dan jauh, meningkatkan kinerja deteksi di berbagai rentang.

Metode berdasarkan proyeksi kamera-LiDAR tradisional sering menggunakan cara penggabungan fitur sederhana seperti penggabungan atau perataan, dengan asumsi pentingnya sama di seluruh jenis data. Namun, gambar RGB dapat mengandung gangguan karena objek yang tertutup atau latar belakang yang berantakan, dan penggabungan naif

dapat mengurangi ketepatan deteksi 3D. Untuk mengatasi hal ini, metode berdasarkan perhatian telah muncul, yang sesuai kebutuhan memberi bobot fitur di seluruh jenis data. Metode penggabungan global menyelaraskan fitur dengan banyak jenis data dalam ruang bersama dan menerapkan perhatian saluran. BEVFusion dan PA3DNet menggunakan perhatian Squeeze-and-Excitation (SE) untuk mengkalibrasi ulang pentingnya saluran.

Namun, perhatian saluran kurang memiliki resolusi ruang. DeepFusion mengatasi hal ini menggunakan cross-attention, memperlakukan fitur LiDAR sebagai query dan fitur gambar sebagai kunci dan nilai, memungkinkan setiap titik 3D untuk fokus pada piksel yang relevan. Namun, metode global mungkin gagal pada objek kecil atau jauh karena informasi redundan di seluruh scene. Metode penggabungan lokal seperti SparseFusion dan SFD mengisolasi instance foreground dan menerapkan attention pada level instance, mengurangi interferensi latar belakang. Namun, mereka mungkin melewatkan isyarat 2D halus selama proposal instance. Pendekatan hibrida menggabungkan penggabungan global dan lokal. 3D-CVF menggunakan channel attention di kedua tahap. AutoAlign dan AutoAlignV2 meningkatkan konsistensi 2D-3D menggunakan cross-attention dan nilai kesalahan kesamaan. LoGoNet melakukan penggabungan paralel fitur scene dan instance, diikuti oleh self-attention untuk interaksi di antara keduanya.

Susunan sistem berbasis Transformer telah menunjukkan kemampuan yang kuat dalam deteksi objek 3D dengan banyak jenis data karena kekuatannya dalam penyesuaian agar selaras dan penggabungan fitur. Kerangka kerja tipikal menggunakan Transformer spesifik jenis data untuk pengkodean intra-modal dan Transformer berdasarkan kueri untuk interaksi lintas-modal. Kueri objek yang dapat dipelajari menyelaraskan penentuan posisi objek di seluruh jenis data dan mengkodekan informasi tingkat instance. CAT-Det mempelopori arah ini dengan memperkenalkan Transformer Dua-aliran dan Lintas-Modal untuk menangkap ketergantungan intra- dan inter-modal jarak jauh. TransFusion menerapkan dekoder Transformer dua tahap, menginisialisasi kotak pembatas dari LiDAR dan menggabungkannya dengan fitur gambar pada tahap kedua. CMT meningkatkan kesadaran ruang dengan menyuntikkan informasi posisi 3D ke dalam potongan data dengan banyak jenis data dan menginisialisasi kueri yang dipandu posisi berdasarkan PETR. Untuk mengurangi false negative akibat objek yang tertutup dan kekacauan, FocalFormer3D memperkenalkan Hard Instance Probing, yang memfokuskan pelatihan pada contoh-contoh sulit dengan secara progresif memperbaiki kandidat objek.

Pembagian objek berdasarkan makna 3D. Metode pembagian objek berdasarkan makna 3D dengan banyak jenis data terkini dapat dikategorikan ke dalam pendekatan berdasarkan penggabungan interaktif, berdasarkan pemindahan pengetahuan, dan berdasarkan penyesuaian bidang tak terawasi (UDA). (1) Metode Berdasarkan Penggabungan Interaktif. Pendekatan ini menggabungkan fitur dari gambar 2D dan kumpulan titik 3D 3D melalui interaksi antarjenis data, seringkali mengikuti struktur U-Net. Meyer dkk. mengusulkan metode pertama semacam itu dengan menyelaraskan titik LiDAR dengan piksel kamera dan memasukkan fitur gabungan ke dalam LaserNet. Karya-karya selanjutnya, seperti Fus-eSeg, meningkatkan pengambilan ciri, sementara PMF memproyeksikan fitur antarjenis data untuk menangani perbedaan modal. RPVNet sesuai kebutuhan menggabungkan pandangan

jangkauan, titik, dan kotak kecil 3D. Lif-Seq memperkenalkan penggabungan kasar-ke-halus dengan rektifikasi offset, dan MSeg3D menggabungkan penggabungan berdasarkan bentuk dan ukuran ruang dan makna dengan perhatian antarmasukan untuk interaksi jarak jauh. (2) Metode Berdasarkan Pemindahan pengetahuan. Karena tumpang tindih bidang pandang (FOV) kamera-LiDAR yang terbatas, metode ini mentransfer pengetahuan 2D ke pengubah masukan 3D selama pelatihan dan membuang masukan 2D pada penarikan kesimpulan. 2DPASS menyelaraskan fitur 3D yang difusikan dan murni melalui pemindahan pengetahuan satu arah. CMDFusion lebih meningkatkan hal ini dengan menyimpan perkiraan awal 2D dalam cabang memori untuk menghasilkan fitur pseudo-2D, memungkinkan penggabungan dua arah selama pelatihan dan penarikan kesimpulan. (3) Metode Berdasarkan UDA. Pelabelan kumpulan titik 3D 3D itu mahal, menjadikan UDA sebagai alternatif yang praktis. Pembelajaran lintas bidang menyelaraskan perkiraan bidang sumber dan target, sementara pembelajaran antarjenis data menyelaraskan perkiraan 2D dan 3D menggunakan perbedaan KL. xMUDA menggabungkan pembelajaran antarjenis data dan pelatihan mandiri, menggunakan label semu untuk memanfaatkan data yang tidak berlabel. MM2D3D memanfaatkan fitur kedalaman dan makna yang kuat di berbagai bidang. Namun, metode sebelumnya sering menyelaraskan fitur satu-ke-satu, mengabaikan fitur yang tidak cocok. DsCML mengatasi hal ini melalui penyaringan pola deformabel, memungkinkan setiap titik 3D untuk mengumpulkan informasi dari beberapa piksel 2D yang relevan melalui pembelajaran sparse-to-dense.

Penandaan Gambar 3D. Metode penandaan gambar 3D (VG) dapat dikategorikan menjadi: pendekatan deteksi-kemudian-cocok dua tahap dan pendekatan panduan bahasa satu tahap. Metode dua tahap pertama-tama menghasilkan proposal melalui detektor atau segmenter 3D yang telah dilatih sebelumnya, kemudian mencocokkan proposal dengan fitur bahasa. Metode awal seperti Scan-Refer menggabungkan fitur untuk pencocokan, sementara ReferIt3D dan TGNN memperkenalkan pembuatan model relasional melalui jaringan grafik. FFL-3DOG menggabungkan grafik adegan dan relasi multi-level, dan InstanceRefer membuat model konteks tingkat instance dan pencocokan fitur holistik. Metode berdasarkan Transformer lebih lanjut meningkatkan penggabungan fitur. LanguageRefer merumuskan 3D VG sebagai tugas pembuatan model bahasa menggunakan BERT. 3DVG-Transformer menggabungkan agregasi sesuai keadaan yang dipandu koordinat dan perhatian multipleks. TransRefer3D mengusulkan perhatian yang sadar entitas dan relasi untuk pengaitan dengan dunia nyata yang lebih detail, sementara MVT menggunakan Transformer multi-view untuk mempelajari gambaran invarian tampilan melalui penggabungan multi-view yang diproyeksikan. Beberapa metode menggabungkan isyarat gambar 2D untuk mengatasi keterbatasan kumpulan titik 3D yang jarang dan bising. SAT menyelaraskan makna 2D (label, fitur, bentuk dan ukuran ruang) dengan proposal 3D menggunakan nilai kesalahan korespondensi objek tambahan. LAR menghasilkan tampilan 2D buatan dari adegan 3D dan menggabungkannya dengan fitur bahasa dan kumpulan titik 3D melalui Transformer dengan banyak jenis data.

Namun, metode dua tahap dibatasi oleh pengambilan sampel keypoint yang tidak relevan dengan bahasa dari detektor 3D, yang menyebabkan target terlewat (dengan proposal

jarang) atau kebingungan karena objek yang berlebihan (dengan proposal padat), seperti yang dicatat dalam. Metode satu tahap menggabungkan deteksi objek dan pengambilan ciri yang dipandu bahasa. 3D-SPS menggunakan pemilihan titik kunci yang peka terhadap bahasa dan penambangan progresif untuk fokus pada wilayah yang relevan. BUTD-DETR, terinspirasi oleh MDETR, menggunakan Transformer yang menggabungkan proposal dan isyarat bahasa. Untuk mencapai granularitas yang lebih halus, EDA memisahkan elemen tekstual (objek, atribut, relasi) dan menyelaraskannya secara individual dengan kandidat 3D, memungkinkan pencocokan yang padat dan dapat diinterpretasikan serta mengurangi ketidakjelasan dari bagian kalimat yang tidak relevan. Keterangan Gambar Padat 3D. Tidak seperti keterangan gambar 2D yang menggambarkan konten gambar global, keterangan gambar padat 3D membutuhkan pemahaman yang lebih baik tentang atribut objek dan lokasi ruang. Metode yang ada terbagi menjadi pendekatan deteksi-kemudian-deskripsi dua tahap dan pendekatan deteksi-deskripsi paralel satu tahap. Metode dua tahap mendeteksi objek terlebih dahulu dan kemudian menghasilkan keterangan per proposal. Misalnya, Scan2Cap membangun grafik relasional di atas proposal objek dan menggunakan perhatian yang sadar konteks untuk keterangan. MORE membuat model relasi rumit dengan mengkodekan relasi ruang orde pertama dan menyusunnya menjadi grafik triplet untuk menyimpulkan makna multi-orde. SpaCap3D secara jelas mengawasi pembelajaran relasional di antara proposal untuk meningkatkan pembuatan model ruang. Untuk memperkaya gambaran makna, beberapa metode menggabungkan fitur gambar 2D.

Meskipun efektif, ini meningkatkan biaya penarikan kesimpulan. X-Trans2Cap mengatasi hal ini dengan pemindahan pengetahuan antarjenis data: jaringan guru belajar dari data 2D dan 3D, membimbing jaringan siswa yang hanya menggunakan masukan 3D selama penarikan kesimpulan. Ini mempertahankan keterangan berkualitas tinggi sambil meningkatkan efisiensi. Metode satu tahap mengatasi keterbatasan dua tahap seperti ketergantungan pada kualitas detektor dan kurangnya perbaikan bersama. Terinspirasi oleh 3DETR, pendekatan ini memperlakukan keterangan padat sebagai masalah perkiraan set-ke-set. Vote2Cap-DETR memperkenalkan kueri suara yang dapat dipelajari dan keterangan dengan petunjuk ganda. Setiap kueri objek mengkodekan kecenderungan yang tidak seimbang ruang dan konteks makna. Dekoder secara bersamaan menghasilkan kotak objek dan keterangan, di mana pembuatan keterangan dipandu oleh fitur tingkat objek dan konteks sekitarnya. Bersama-sama, pendekatan ini mencerminkan pergeseran menuju Memahami Ruang 3D yang terpadu, efisien, dan sadar konteks untuk keterangan padat.

Penjawaban Pertanyaan 3D. 3D Gambar Question Answering (VQA) membutuhkan pemahaman ruang dan proses berpikir yang kuat atas adegan yang rumit. ScanQA adalah pendekatan VQA 3D representatif yang memperkirakan jawaban beserta kotak pembatas untuk objek yang dirujuk dalam pertanyaan. Pertama, ia mengkodekan proposal objek dan kata-kata pertanyaan menggunakan pengubah masukan gambar dan bahasa, kemudian menggabungkannya melalui lapisan Transformer. Fitur yang digabungkan digunakan untuk memperkirakan lokasi objek, label, dan jawaban, dengan nilai kesalahan deteksi tambahan yang meningkatkan pemahaman ruang. 3D Situated Question Answering (SQA) memperluas

VQA 3D dengan menggabungkan informasi sistem AI orang pertama. Diberikan adegan 3D, pertanyaan, dan deskripsi situasi (misalnya, posisi dan orientasi sistem AI), model harus menjawab dari perspektif sistem AI. SQA3D memperkenalkan dua baseline: model ScanQA yang dimodifikasi dengan tambahan head perkiraan lokasi, dan model tanpa contoh khusus yang memanfaatkan Model AI Bahasa Besar (LLM). Yang terakhir menggunakan sistem keterangan 3D untuk mengubah adegan menjadi deskripsi tekstual, yang, dikombinasikan dengan konteks situasi, dimasukkan ke dalam LLM seperti GPT-3 atau Unified QA untuk menghasilkan jawaban tanpa pelatihan khusus tugas.

Generasi Adegan 3D Berdasarkan Teks. Generasi adegan 3D berdasarkan teks bertujuan untuk mensintesis lingkungan 3D realistis dari deskripsi bahasa alami. Karena mengumpulkan pasangan adegan-teks skala besar tidak praktis, banyak metode memanfaatkan model difusi teks-ke-gambar 2D yang telah dilatih sebelumnya untuk pengawasan. Metode berdasarkan mesh, seperti SceneScape dan Text2Room, menghasilkan adegan 3D dengan memperbarui gambaran mesh secara berulang. Mereka menggabungkan prioritas konten dari model difusi dengan prioritas berdasarkan bentuk dan ukuran dari estimator kedalaman satu kamera untuk membangun dan memperbaiki mesh dari beberapa tampilan. Namun, penerapannya sebagian besar terbatas pada adegan dalam ruangan karena keterbatasan mesh dan kerumitan berdasarkan bentuk dan ukuran lingkungan luar ruangan.

Metode berdasarkan *Neural Radiance Field* (NeRF) menawarkan detail yang lebih halus dan artefak yang lebih sedikit. DreamFusion memperbaiki NeRF menggunakan Score Distillation Sampling untuk menyelaraskan pembuatan gambar 2D dengan model difusi. GAUDI mengadopsi kerangka kerja dua tahap: memisahkan pose kamera dan gambaran adegan sebelum pembuatan model yang menghasilkan data melalui DDPM. Generasi komposisi ditangani oleh Comd3D, yang mengedit sub-wilayah VoxNeRF melalui model difusi spesifik wilayah yang dipandu oleh bounding box dan petunjuk teks. Untuk meningkatkan modularitas dan penggunaan kembali, Set-the-Scene memperkenalkan kumpulan NeRF independen sebagai proksi objek. Ia bergantian antara perbaikan lokal objek individual dan perbaikan adegan global. Menerapkan transformasi kaku untuk menyelaraskan setiap objek dalam adegan penuh memungkinkan konstruksi adegan komposisional dan dapat diedit dengan konsistensi gaya.

Pengenalan 3D Kosakata Terbuka. Patokan uji pengenalan 3D saat ini bergantung pada label penjelas manusia dengan makna kumpulan tertutup, seperti 20 kelas dalam Scan-Net, yang membatasi kemampuan model terawasi untuk menangani kategori kosakata terbuka. Model dasar penglihatan-bahasa seperti CLIP, yang dilatih pada pasangan gambar-teks skala besar, menunjukkan proses berpikir kosakata terbuka yang kuat dan telah menginspirasi cara pengenalan 3D baru. CLIP2Scene mentransfer pengetahuan CLIP ke jaringan 3D melalui pembatasan makna. Pertama, ia menghasilkan pasangan piksel-teks padat menggunakan MaskCLIP, memetakannya ke kumpulan titik 3D 3D melalui penyesuaian multi-view, dan memilih sampel titik kontras berdasarkan makna teks untuk pelatihan. RegionPLC lebih lanjut meningkatkan pemahaman 3D melalui pembelajaran kontras bahasa titik regional. Ia menghasilkan petunjuk dan keterangan 2D yang padat, menyelaraskannya dengan wilayah 3D

menggunakan bentuk dan ukuran ruang, dan mempelajari bentuk angka ringkas tingkat titik diskriminatif dengan pengurangan gangguan dari tetangga yang tidak relevan. Untuk menjembatani jenis data 2D dan 3D, metode seperti OpenScene menyelaraskan gambar RGB-D dengan kumpulan titik 3D menggunakan pembelajaran dengan membandingkan yang dipandu CLIP.

Alih-alih mengandalkan penyesuaian agar selaras global, CLIP-FO3D memperkenalkan potongan data klasifikasi lokal dalam CLIP untuk mengekstrak fitur gambar regional, yang kemudian diproyeksikan ke ruang 3D. Model 3D dilatih untuk membuat sekecil mungkin jarak antara fitur yang diproyeksikan ini dan fitur titik yang dipelajari, memungkinkan kaitan dengan makna yang sadar wilayah. Pendekatan ini memanfaatkan pengetahuan CLIP untuk memungkinkan pengenalan 3D kosakata terbuka dengan menyelaraskan fitur kumpulan titik 3D dengan makna bahasa melalui perantara gambar 2D. Pelatihan Awal Penglihatan-Bahasa 3D. Pendekatan bahasa-penglihatan 3D saat ini seringkali spesifik untuk tugas tertentu dan tidak memiliki cara kerja yang menyatu untuk gambaran antarjenis data tujuan umum. Untuk mengatasi hal ini, 3D-VisTA mengusulkan Transformer yang telah dilatih sebelumnya untuk penyesuaian agar selaras teks-adegan 3D, yang dapat disesuaikan dengan tugas-tugas seperti penentuan dasar gambar, pembuatan keterangan padat, dan menjawab pertanyaan. Ia menggunakan self-attention untuk pengkodean modal tunggal dan penggabungan dengan banyak jenis data, tanpa kustomisasi khusus tugas.

Untuk mendukung pelatihan awal, 3D-VisTA memperkenalkan ScanScribe, kumpulan data skala besar pertama dari pasangan teks-adegan 3D. Dengan menggunakan pembuatan model bahasa/objek bertopeng dan pencocokan teks-adegan, ia mempelajari penyesuaian agar selaras teks-3D yang efektif. Demikian pula, 3D-VLP meningkatkan pemahaman antarjenis data melalui penyesuaian agar selaras makna ruang yang sadar konteks dan pembuatan model bertopeng bahasa-3D timbal balik. Ia menyelaraskan objek 3D dengan kata benda kelas dan ekspresi referensi, menangkap hubungan ruang. Perkiraan bersama label objek bertopeng dan teks bertopeng memungkinkan pertukaran makna antara bentuk dan ukuran ruang 3D dan bahasa.

Penjelajahan Lingkungan Secara Aktif

Eksplorasi aktif mengacu pada kemampuan sistem AI untuk memperoleh informasi, membangun gambaran di dalam sistem, dan menyelesaikan tujuan tugas melalui pengambilan keputusan mandiri dan perilaku dapat menyesuaikan diri di lingkungan yang tidak dikenal atau sebagian dikenal. Tidak seperti pengamatan pasif tradisional, eksplorasi aktif menekankan bahwa AI tidak hanya harus mempersepsikan lingkungannya tetapi juga memahami keadaannya sendiri dan menggunakan pemahaman ini untuk memandu tindakan di masa mendatang. Kemampuan ini mendasar untuk mencapai kecerdasan seperti manusia, karena manusia menyesuaikan diri dengan lingkungan baru, memecahkan masalah, dan mengumpulkan pengalaman melalui eksplorasi berkelanjutan. Dalam evolusi kecerdasan buatan, pentingnya eksplorasi aktif melampaui inovasi teknis. Ini menandai pergeseran dari pengenalan tetap ke proses berpikir berubah-ubah, memungkinkan mesin untuk beroperasi dengan otonomi dan kemampuan menyesuaikan diri yang lebih besar dalam keadaan dunia

nyata yang rumit dan selalu berubah. Misalnya, dalam bidang seperti navigasi robot, interaksi asisten virtual, dan mengemudi mandiri, kemampuan untuk eksplorasi aktif secara langsung berdampak pada efisiensi dan keandalan sistem.

Eksplorasi aktif berfokus pada memungkinkan sistem AI untuk secara mandiri memperoleh pengamatan informatif melalui interaksi dengan lingkungan. Pendekatan awal bergantung pada pembangunan model lingkungan jelas atau tidak langsung. Metode berdasarkan model yang representatif meliputi MAX dan Active Neural SLAM, yang memanfaatkan pembuatan model dan pemetaan untuk memperkirakan untuk mendukung navigasi yang efisien di ruang yang tidak terlihat. Sebaliknya, APT dan DBMF-BPI berfokus pada eksplorasi tanpa model melalui interaksi lingkungan langsung untuk mengurangi ketergantungan pada pembuatan model jelas. Upaya terbaru lebih meningkatkan kemampuan eksplorasi dengan menggabungkan pengamatan dengan banyak jenis data dan proses berpikir berdasarkan makna.

Proses Berpikir Berwujud

Proses Berpikir Berwujud mengacu pada munculnya gambaran di dalam sistem dan kemampuan proses berpikir selama interaksi, yang didorong oleh refleksi diri sistem AI terhadap pengamatan dan pengalaman yang terakumulasi. Bagian ini membentuk inti dari AI berwujud, memungkinkan sistem AI untuk melakukan perencanaan tugas, penarikan kesimpulan sebab-akibat, dan proses berpikir jangka panjang. Studi terbaru tentang Proses Berpikir Berwujud terutama berfokus pada tiga aspek: perencanaan diri yang didorong oleh tugas, refleksi diri yang didorong oleh memori, dan model dasar dengan banyak jenis data berwujud. Tabel 1.2 menyajikan metode representatif yang dianalisis dari empat perspektif: jenis data masukan, tipe proses berpikir, mode proses berpikir, dan tipe keluaran. Dimensi-dimensi ini mencerminkan bagaimana sistem AI berwujud mempersepsikan informasi, membentuk model internal, dan melakukan proses berpikir.

Perencanaan Diri yang Didorong oleh Tugas. Dalam perencanaan mandiri berdasarkan tugas, sistem AI secara mandiri menghasilkan rencana terstruktur berdasarkan tujuan tugas, konteks lingkungan, dan pengetahuan internal, tanpa instruksi manusia yang jelas. Pembelajaran terstruktur adalah solusi klasik yang mengembangkan ruang perencanaan tersembunyi atau pemetaan strategi langsung, mencapai efisiensi tinggi dalam distribusi pelatihan tetapi kurang tangguh terhadap keadaan di luar distribusi. Pendekatan representatif meliputi L3P, Egoplaner, dan ETPNav. Kemajuan terbaru menggabungkan LLM atau model yang menghasilkan data ke dalam perencanaan mandiri. LLM-Planner dan AutoAct menggabungkan LLM ke dalam perencanaan dengan mengaitkan proses berpikir yang dipandu bahasa ke dalam berbagai tugas, sementara RPG menawarkan perspektif yang menghasilkan data, bertujuan untuk menyatukan perencanaan dan pembuatan konten melalui proses berpikir dengan banyak jenis data.

Tabel 1.2 Perbandingan proses berpikir berwujud (I: Citra, L: Bahasa, C: Kumpulan titik tiga dimensi)

Metode	Tahun	Jenis data	Jenis	Mode proses berpikir	Keluaran
Perencanaan mandiri berdasarkan tugas					
L3P	2021	I + L	Perencana	Neural + Simbolik	Tindakan
LLM-Planner	2023	I + L	Perencana	Neural + Simbolik	Tindakan
Egoplaner	2023	I	Perencana	Simbolik	Tindakan
AutoAct	2024	L	Perencana	Neural	Tindakan
RPG	2024	I + L	Perencana	Neural	Strategi
EIPNav	2024	I + L	Perencana	Neural + Simbolik	Strategi
Refleksi mandiri berdasarkan memori					
Reflexion	2023	L	Memori	Beam + Replay	Strategi
Reflect	2023	I + L	Memori	Neural + Simbolik	Strategi
RILA	2024	L	Memori	Neural	Strategi
Optimus-1	2024	I + L	Memori	Neural	Strategi
EvoAgent	2025	I + L	Memori	Neural	Strategi
REMAC	2025	L	Memori	Neural + Simbolik	Strategi
Model dasar dengan banyak jenis data berwujud					
SayCan	2022	I + L	Perencana + Penyelaras	Neural	Jawaban + Tindakan
GATO	2022	I + L + P	Penyelaras	Neural	Tindakan
EmbodiedGPT	2023	I + L	Penyelaras	Neural	Jawaban + Tindakan
Kosmos-2	2023	I + L	Penyelaras	Neural	Jawaban
MultiPLY	2024	I + L	Penyelaras	Neural	Jawaban
ManipLLM	2024	I + L	Penyelaras	Neural	Jawaban + Tindakan

Refleksi Diri Berdasarkan Memori. Refleksi diri berdasarkan memori memungkinkan sistem AI untuk memanfaatkan pengalaman masa lalu untuk proses berpikir jangka panjang, koreksi kesalahan, dan peningkatan diri. Studi awal berfokus pada pemrosesan memori, termasuk tempat penyimpanan sementara replay ukuran tetap dan susunan sistem memori yang dapat dibedakan. Kemajuan terbaru memperkenalkan cara kerja reflektif, di mana sistem AI meringkas atau mengungkapkan pengalaman masa lalu untuk memandu keputusan di masa depan. Reflexion dan Reflect memungkinkan sistem AI untuk secara berulang mengoreksi diri

sendiri dengan menggabungkan umpan balik verbal ke dalam perencanaan tindakan, sementara RILA memperluas proses berpikir reflektif ke navigasi makna dengan banyak jenis data. Di luar refleksi individu, Optimus-1 dan REMAC menggabungkan memori dengan banyak jenis data atau banyak sistem AI untuk mendukung kerja sama jangka panjang. EvoAgent lebih lanjut memajukan arah ini dengan menggabungkan pembuatan model dunia berkelanjutan dengan perencana berdasarkan memori, memungkinkan evolusi mandiri sepenuhnya di seluruh tugas berurutan.

Model Dasar Dengan banyak jenis data Berwujud. Di era MLLM, model dasar dengan banyak jenis data berwujud telah muncul sebagai salah satu solusi paling menjanjikan untuk menyatukan perencanaan, proses berpikir, dan kemampuan terkait proses berpikir berwujud lainnya. Kemajuan terkini didorong oleh konstruksi data dan pengembangan model. Upaya data berfokus pada pembangunan patokan uji berkualitas tinggi untuk mendukung penilaian yang mudah dikembangkan dan bermakna secara terkait proses berpikir, seperti MuEP, ECBench, MFE-ETP, dan EmbodiedBench. Dari sisi model, kemajuan terbaru mencakup sistem AI berdasarkan kegunaan benda (misalnya, SayCan dan GATO) yang menyelaraskan pemahaman bahasa dengan pilihan tindakan berwujud, pendekatan pelatihan awal bahasa-penglihatan (seperti EmbodiedGPT dan Kosmos-2) yang mempromosikan proses berpikir berwujud yang mudah dikembangkan, dan desain berpusat pada objek (seperti MultiPLY dan ManipLLM) yang meningkatkan kemampuan penggerakan benda dan interaksi. Model-model ini secara bersama-sama bertujuan untuk membangun AI berwujud yang dapat ditransfer dan digeneralisasi.

Interaksi yang Menyesuaikan

Interaksi yang Menyesuaikan mengacu pada proses di mana sistem AI memengaruhi lingkungan melalui tindakan atau perilaku yang didasarkan pada pengamatan dan kognisinya. Riset yang ada menekankan pentingnya kemampuan ini dalam memungkinkan sistem AI tidak hanya untuk merespons tetapi juga untuk mengubah lingkungan sekitarnya. Studi tentang Interaksi yang Menyesuaikan mencakup kendali tindakan, interaksi perilaku, dan pengambilan keputusan bekerja sama. Untuk lebih memahami metode yang ada, kami menganalisis pendekatan representatif dari empat perspektif, termasuk jenis data masukan, jenis interaksi, cara pembelajaran, dan jenis tugas, seperti yang ditunjukkan pada Tabel 1.3. Dimensi-dimensi ini mencerminkan bagaimana sistem AI merasakan lingkungan, menentukan tingkat dan struktur interaksi, dan menghasilkan perilaku yang sesuai dalam keadaan banyak sistem AI berubah-ubah atau manusia dalam lingkaran.

Tabel 1.3 Perbandingan Interaksi Menyesuaikan: kendali tindakan, interaksi perilaku, & keputusan bersama berdasar masukan, interaksi, model, & tugas (I: Citra, L: Bahasa, S: Keadaan, P: Rasa posisi tubuh, T: Jalur gerak, IL: Belajar meniru).

Metode	Tahun	Jenis data masukan	Jenis interaksi	Cara pembelajaran	Jenis tugas
Kendali tindakan					

MineDojo	2022	I + L	Perencanaan tingkat tinggi	LLM	Mengikuti instruksi
PaLM-E	2023	I + L + P	Kendali tingkat rendah	MLLM	Penggerakan benda berwujud
RT-2	2023	I + L	Kendali tingkat rendah	VLA	Penggerakan benda berwujud
OpenVLA	2024	I + L	Kendali tingkat rendah	VLA	Penggerakan benda berwujud
Cogagent	2024	I + L	Kendali tingkat rendah	MLLM	Mengikuti instruksi
Octo	2024	I + L + P	Kendali tingkat rendah	VLA	Penggerakan benda berwujud
CrossFormer	2024	I + L	Kendali tingkat rendah	VLA	Penggerakan benda berwujud
HPT	2024	I + L	Kendali tingkat rendah	VLA	Penggerakan benda berwujud
Interaksi perilaku					
GAIL	2016	T	Perilaku	IL	Pembelajaran lintasan
MGAIL	2017	T	Perilaku	IL	Pembelajaran lintasan
TrafficSim	2021	T	Perilaku	RL	Pembelajaran lintasan
TrajGen	2022	I + T	Perilaku	RL	Pembelajaran lintasan
Behavior-1K	2023	I	Lintasan	IL	Pemahaman perilaku
AgentLens	2024	I + S	Lintasan	IL	Pemahaman perilaku
ECL	2024	I + L	Perencanaan tingkat tinggi	IL	Penggerakan benda berwujud
Keputusan bekerja sama					
QMIX	2018	S	Perilaku	RL	Keputusan kooperatif
Qtran	2019	S	Perilaku	RL	Keputusan kooperatif
QPLEX	2019	S	Perilaku	RL	Keputusan kooperatif
MAT	2022	S	Perilaku	RL	Keputusan kooperatif

CoELA	2024	I + L	Kendali tingkat rendah	LLM	Penggerakan benda kooperatif
AgentVerse	2024	L	Perencanaan tingkat tinggi	LLM	Simulasi masyarakat sistem AI
MetaGPT	2024	L	Perencanaan tingkat tinggi	LLM	Simulasi masyarakat sistem AI
Combo	2024	L	Perencanaan tingkat tinggi	LLM	Perencanaan kooperatif

Kendali Tindakan. Kendali tindakan menghasilkan perintah motorik untuk interaksi berwujud. Metode awal didasarkan pada teori kendali dengan pembuatan model sistem berubah-ubah atau RL melalui coba-coba. Yang pertama efektif untuk tugas terstruktur atau berulang, sedangkan yang terakhir dapat diadaptasi untuk masalah nonlinier memiliki banyak unsur. Kemajuan terbaru terutama mengikuti tiga arah. Model penglihatan-bahasa-tindakan (VLA), seperti PaLM-E, RT-2, OpenVLA, dan CogAgent, menggabungkan proses berpikir yang dipandu bahasa untuk kendali yang luwes dan telah ditinjau secara menyeluruh oleh Ma dan kawan-kawan.. Kerangka kerja terbuka seperti MineDojo mendorong akuisisi keterampilan berkelanjutan dari pengetahuan dunia terbuka. Selain itu, pembelajaran lintas bentuk fisik, termasuk CrossFormer, HPT, dan Octo, bertujuan untuk menyatukan pembelajaran strategi di berbagai robot dan jenis data.

Interaksi Perilaku. Perilaku sistem AI terdiri dari serangkaian tindakan. Dibandingkan dengan kendali tindakan, ini menekankan kendali tingkat tinggi melalui pola tindakan yang bermakna, memungkinkan sistem AI untuk berinteraksi secara luwes dan terarah pada tujuan. Kemajuan terbaru terutama terbagi menjadi dua arah. Belajar dengan meniru, termasuk GAIL, MGAIL, TrafficSim, dan TrajGen, memungkinkan akuisisi dan simulasi perilaku rumit secara efisien. BEHAVIOR-1K menyediakan patokan uji skala besar untuk menilai kemampuan berlaku pada keadaan lain perilaku di 1.000 tugas berwujud. Metode peningkatan yang sadar perilaku, seperti AgentLens dan ECL, meningkatkan kekokohan dan penafsiran strategi. Terlepas dari kemajuan ini, mencapai interaksi perilaku jangka panjang yang andal di bawah umpan balik yang jarang tetap menantang.

Pengambilan Keputusan Bekerja sama. Pengambilan keputusan bekerja sama berfokus pada koordinasi beberapa sistem AI untuk mencapai tujuan bersama, yang penting untuk sistem banyak sistem AI dan kerja sama manusia-robot. RL banyak sistem AI adalah solusi klasik, dengan metode seperti QTRAN, QPLEX, dan Qatten yang menangani kerja sama melalui pelatihan dipusatkan dengan pelaksanaan dibagi ke tiap bagian. MAT merumuskan kembali MARL sebagai masalah pembuatan model urutan untuk mengurangi keterbatasan kemampuan berkembang dalam RL banyak sistem AI. Kemajuan terbaru menggabungkan LLM dan model dunia untuk meningkatkan kerja sama banyak sistem AI. MetaGPT, CoELA, dan AgentVerse memanfaatkan LLM untuk proses berpikir dan koordinasi tugas, sementara

COMBO menyusun model dunia terbagi menjadi beberapa bagian untuk mendukung pengambilan keputusan bekerja sama yang mudah dikembangkan.

Bentuk Perangkat Keras

Seiring evolusi AI berwujud, kerumitan dan ukuran model telah meningkat, sehingga meningkatkan kebutuhan perhitungan komputer dan energi. Sistem berwujud, yang sering beroperasi di lingkungan dunia nyata yang berubah-ubah, menghadapi batasan waktu tunda dan daya yang ketat, terutama di edge. Dengan demikian, mengembangkan arah yang ramah perangkat keras yang mempertahankan kinerja sambil memperbaiki efisiensi sangat penting untuk memungkinkan sistem AI berwujud yang responsif dan hemat energi. Perbaikan perangkat keras dalam AI berwujud biasanya mencakup empat bagian: peringkasan model yang sadar perangkat keras, perbaikan tingkat kompiler, akselerator khusus bidang, dan desain bersama perangkat keras-perangkat lunak.

Peringkasan Model yang Sadar Perangkat Keras. Kuantisasi dan pemangkasan adalah cara kunci untuk mengurangi ukuran model dan biaya perhitungan komputer. Pada sistem AI berwujud, yang sering berjalan pada perangkat keras tertanam berdaya rendah, cara tersebut sangat penting untuk memungkinkan penarikan kesimpulan yang cepat dan efisien. Kuantisasi memetakan bobot dan pengaktifan ke lebar bit yang lebih rendah, sementara pemangkasan menghilangkan nilai pengatur yang berlebihan. Untuk mendukung tugas-tugas berwujud di dunia nyata, seperti kendali robot atau navigasi gambar, ukuran penilaian efisiensi perangkat keras seperti daya, kinerja, dan area (PPA) dapat memandu alokasi lebar bit atau rasio pemangkasan, memungkinkan pertukaran spesifik tugas antara ketepatan dan kemampuan penggunaan pada platform fisik.

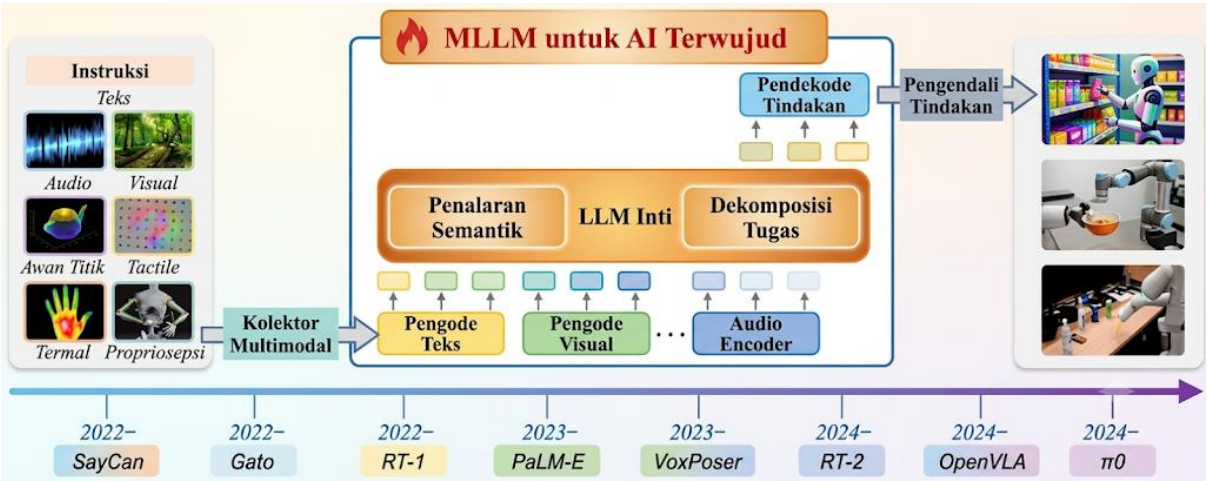
Perbaikan Tingkat Kompiler. Kompiler menjembatani model AI berwujud tingkat tinggi dan pelaksanaan perangkat keras. Dalam sistem berwujud secara real-time, toolchain kompilator sangat penting untuk pemrosesan data sensor dan pengambilan keputusan yang efisien. TVM, yang dibangun di atas LLVM dan CUDA, menghasilkan kode yang diperbaiki di berbagai platform. Kompiler ini mengubah grafik perhitungan komputer melalui penggabungan operator dan penghapusan perhitungan komputer yang berlebihan, memungkinkan perilaku responsif. Strategi pemetaan seperti penataan ulang loop dan tiling meningkatkan lokalitas data, paralelisme, dan akses memori, yang semuanya penting untuk mempertahankan penarikan kesimpulan waktu tunda rendah pada sistem AI berwujud. Akselerator Khusus Bidang. Dengan meningkatnya tuntutan perhitungan komputer, akselerator khusus bidang (DSA) merupakan solusi yang menjanjikan untuk AI berwujud. Sistem ini, dari robot hingga sistem AI AR/VR, mendapat manfaat dari perangkat keras yang cepat dan hemat energi yang dirancang untuk operasi yang sering. TPU Google, yang biasanya tergabung dengan CPU dan GPU melalui PCIe, mempercepat operasi kunci seperti perkalian matriks. Akselerator berdasarkan FPGA memungkinkan susunan ulang untuk menyesuaikan diri dengan tugas baru atau beban kerja yang berubah; akselerator CGRA meningkatkan perhitungan komputer terstruktur dan berat aliran data yang umum dalam pengamatan atau kendali. Sementara itu, akselerator berdasarkan ASIC menawarkan throughput tinggi dan

efisiensi energi, ideal untuk menerapkan model berwujud berkinerja tinggi di lingkungan dunia nyata.

Desain Bersama Perangkat Keras-Perangkat Lunak. Memisahkan desain langkah kerja dan perangkat keras dapat menurunkan efisiensi waktu pelaksanaan. Desain bersama perangkat keras-perangkat lunak mengatasi hal ini melalui perbaikan bersama langkah kerja-sistem dan langkah kerja-perangkat keras. Perbaikan bersama langkah kerja-sistem berfokus pada bagaimana memanfaatkan sepenuhnya sumber daya GPU seperti inti tensor dan inti CUDA untuk mendukung langkah kerja dengan lebih baik. Perbaikan bersama langkah kerja-perangkat keras bertujuan untuk meningkatkan efisiensi penggunaan dengan menyetel model dan susunan sistem perangkat keras. Misalnya, kita dapat melakukan perbaikan multi-objektif berdasarkan jenis operator dalam jaringan dan nilai pengatur susunan perangkat keras. Kita juga dapat merancang skema kuantisasi numerik yang berbeda bersama dengan akselerator perangkat keras yang sesuai untuk mendukung tugas AI berwujud dengan lebih baik.

1.4 ARSITEKTUR UTAMA AI BERWUJUD

Susunan sistem AI berwujud mencerminkan strategi berbeda untuk menggabungkan pengamatan, proses berpikir, dan tindakan dalam cara kerja yang menyatu. Sementara desain tradisional sering menekankan pemisahan terbagi menjadi beberapa bagian, riset terbaru menekankan perlunya susunan sistem ujung-ke-ujung yang mampu menjembatani pemahaman makna tingkat umum dengan interaksi fisik yang mendasar. Bagian ini meninjau tiga arah representatif: LLM/MLLM untuk AI berwujud, model dunia untuk AI berwujud, dan penggabungan MLLM dan model dunia untuk AI berwujud. Dengan memeriksa pendekatan-pendekatan ini, kami mengilustrasikan bagaimana desain susunan sistem membentuk keseimbangan antara proses berpikir berdasarkan makna dan perubahan yang sesuai dengan fisika, menawarkan wawasan tentang konstruksi sistem AI berwujud yang dapat diskalakan dan dapat menyesuaikan diri.



Gambar 1.3 Peta jalan pengembangan MLLM untuk AI berwujud

LLM Meningkatkan AI berwujud

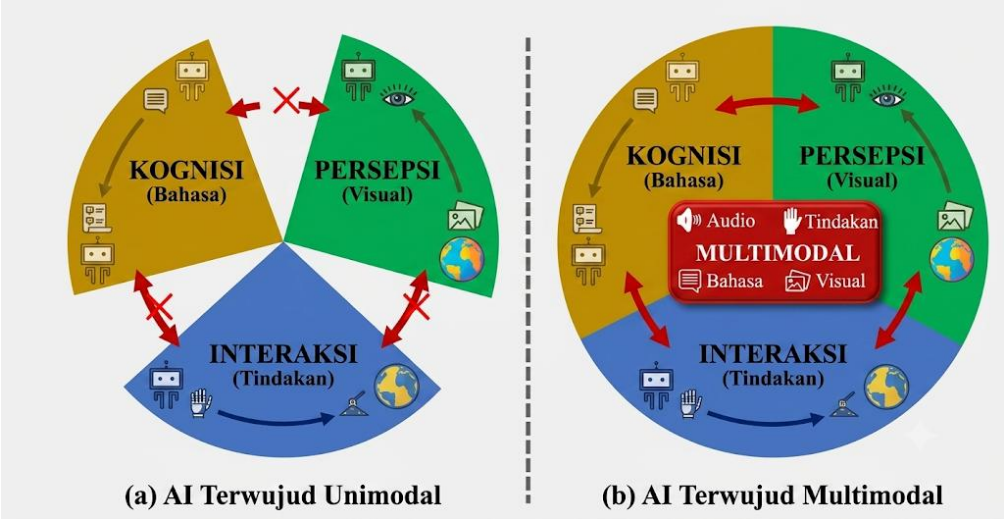
LLM memberdayakan AI berwujud melalui proses berpikir berdasarkan makna dan pemecahan tugas, membawa instruksi bahasa alami tingkat tinggi dan tindakan bahasa alami tingkat rendah ke dalam Proses Berpikir Berwujud. Proses berpikir berdasarkan makna. Proses berpikir berdasarkan makna bergantung pada LLM untuk mengekstrak dan memahami makna dari instruksi bahasa alami dengan memeriksa pola dalam teks, menghubungkan kata-kata dengan konteksnya, dan memanfaatkan pengetahuan sebelumnya. Dengan menggunakan struktur transformer, model-model ini menerjemahkan kata-kata masukan ke dalam gambaran di dalam sistem yang menangkap berbagai lapisan makna, dari sintaksis hingga konteks dunia nyata. Bagian sistem perhatian membantu memfokuskan pada informasi yang berguna sambil menyaring kebisingan yang tidak relevan, mendukung penarikan kesimpulan dan pembuatan analogi. Dengan pengetahuan yang mereka peroleh selama pelatihan awal skala besar, LLM menggabungkan informasi dunia umum dengan masukan tugas spesifik untuk menyusun struktur berdasarkan gagasan yang menghubungkan apa yang ditulis dengan apa yang perlu dilakukan. Cara kerja ini memungkinkan LLM untuk melakukan proses berpikir langkah demi langkah dengan memperkirakan kemungkinan potongan data berikutnya dengan cara yang memeriksa konsistensi dan memastikan makna yang dimaksud tetap jelas.

Pemecahan tugas. Pemecahan tugas menggunakan logika berurutan LLM untuk memecah tujuan rumit menjadi tugas kecil dengan menganalisis secara bertingkat ketergantungan sesuai keadaan dan keselarasan tujuan. Dengan memanfaatkan pemicu rantai pemikiran, LLM secara berulang memecah instruksi menjadi langkah-langkah yang dapat ditindaklanjuti, memprioritaskan saling ketergantungan sambil menyelesaikan ketidakjelasan melalui pemeriksaan keselarasan makna. Karya-karya representatif seperti SayCan pertama-tama menyediakan kumpulan tindakan bahasa alami yang telah dilatih untuk dunia nyata, yang digunakan untuk membatasi LLM agar mengusulkan tindakan yang layak dan sesuai konteks; kemudian menggunakan LLM untuk mengubah instruksi bahasa alami menjadi urutan tindakan bahasa alami; akhirnya menggunakan rumus nilai untuk memeriksa kelayakan urutan tindakan bahasa alami dalam lingkungan fisik tertentu. Karya-karya ini menunjukkan bahwa LLM sangat berguna bagi robot yang bertujuan untuk bertindak berdasarkan instruksi tingkat tinggi yang berlaku dalam jangka waktu tertentu yang diungkapkan dalam bahasa alami. Namun, LLM hanyalah bagian dari keseluruhan sistem AI berwujud, yang dibatasi oleh kumpulan tindakan bahasa alami yang tetap dan lingkungan fisik tertentu, dan sulit untuk mencapai perluasan dapat menyesuaikan diri pada robot dan lingkungan baru.

Dari Dengan satu jenis data ke Dengan banyak jenis data

Perkembangan AI berwujud telah berkembang dari sistem dengan satu jenis data ke dengan banyak jenis data, seperti yang ditunjukkan pada Gambar 1.4. Awalnya, AI berwujud terutama berkaitan dengan jenis data individu, seperti penglihatan, bahasa, atau tindakan, di mana pengamatan, proses berpikir, dan interaksi didorong oleh satu masukan dari sensor. Seiring perkembangan bidang ini, keterbatasan AI berwujud dengan satu jenis data menjadi jelas, dan telah terjadi pergeseran besar menuju pengintegrasian berbagai jenis data dari sensor. AI berwujud dengan banyak jenis data sekarang dipandang penting untuk menciptakan

sistem AI yang lebih dapat menyesuaikan diri, luwes, dan tangguh yang mampu melakukan tugas-tugas rumit di lingkungan yang berubah-ubah.



Gambar 1.4 Perbandingan AI berwujud dengan satu jenis data dan banyak jenis data

AI berwujud dengan satu jenis data telah mendapat manfaat dari perkembangan pesat di bidang-bidang dasar seperti penglihatan komputer, pemrosesan bahasa alami, dan Belajar melalui imbalan. Metode dengan satu jenis data ini unggul dalam menangani bagian sistem spesifik dalam AI berwujud. Misalnya, cara penglihatan komputer telah mendorong kemajuan dalam SLAM gambar dan Memahami Ruang 3D dalam bagian sistem Pengamatan Aktif. Cara pemrosesan bahasa alami, terutama LLM, telah menjadi solusi populer untuk mengatasi perencanaan tugas dan proses berpikir jangka panjang dalam bagian sistem Proses Berpikir Berwujud. Meskipun AI berwujud dengan satu jenis data berkinerja baik dalam bagian sistem independen, ia selalu menghadapi dua keterbatasan bawaan. Di satu sisi, informasi yang terkandung dalam satu jenis data terbatas, menghambat kinerja pengamatan, proses berpikir, dan interaksi. Misalnya, sistem gambar saja kesulitan memahami lingkungan dalam pengaturan berubah-ubah atau ambigu, sementara sistem berdasarkan pendengaran menghadapi tantangan dalam kebisingan dunia nyata dan pemrosesan sinyal. Di sisi lain, jenis data yang beragam dan beragam menghambat transfer dan berbagi informasi antar bagian sistem. Pengamatan sistem AI terhadap lingkungan gagal memudahkan pembentukan kognisinya, sementara evolusi proses berpikir gagal memudahkan interaksi dengan lingkungan.

Sebaliknya, AI berwujud secara dengan banyak jenis data telah muncul sebagai cara yang lebih menjanjikan. Dengan menggabungkan data dari berbagai jenis data penginderaan, seperti umpan balik gambar, pendengaran, dan penciuman, metode ini dapat memberikan pemahaman yang lebih holistik dan tepat tentang lingkungan. Lebih penting lagi, AI berwujud secara dengan banyak jenis data dapat memudahkan penggabungan yang lebih dalam antara pengamatan, proses berpikir, dan interaksi. Kemajuan terbaru dalam MLLM dan model dunia memungkinkan sistem AI untuk menangani berbagai jenis data secara lebih efektif, menjanjikan peningkatan kemampuan AI berwujud. Penggabungan model-model ini dianggap

sebagai langkah kunci untuk memungkinkan AI berwujud secara dengan banyak jenis data di lingkungan yang berubah-ubah dan tidak pasti.

MLLM Meningkatkan AI berwujud. MLLM dapat bertindak pada seluruh sistem AI berwujud dan dapat menyelesaikan masalah LLM dengan baik dengan menjembatani masukan dengan banyak jenis data tingkat tinggi dan urutan tindakan gerak tingkat rendah ke dalam penggunaan AI berwujud ujung-ke-ujung (seperti yang ditunjukkan pada Gambar 1.3). Dibandingkan dengan LLM, proses berpikir berdasarkan makna memanfaatkan pemahaman antarjenis data MLLM untuk menafsirkan makna dari masukan gambar, suara, atau sentuhan, misalnya, mengenali objek, menyimpulkan hubungan ruang, atau memperkirakan perubahan lingkungan. Bersamaan dengan itu, pemecahan tugas menggunakan logika berurutan MLLM untuk memecah tujuan rumit menjadi tugas kecil sambil terus menyesuaikan rencana berdasarkan umpan balik sensor. MLLM terutama mencakup Model Bahasa-Penglihatan (VLM) dan model Bahasa-Penglihatan-Tindakan (VLA).

VLM untuk AI berwujud. VLM untuk AI berwujud menggabungkan pemahaman instruksi gambar dan bahasa untuk memungkinkan sistem AI fisik atau virtual untuk mempersepsikan lingkungan mereka dalam tugas-tugas yang berfokus pada tujuan. Karya-karya representatif seperti PaLM-E pertama-tama melatih pengkodean gambar dan bahasa secara ujung-ke-ujung, bersamaan dengan Model AI Bahasa Besar yang telah dilatih sebelumnya; kemudian menggabungkan hasil pengkodean jenis data sensor berkelanjutan dunia nyata ke dalam VLM dan membangun hubungan antara kata dan pengamatan; akhirnya, mencapai penyelesaian banyak tugas melalui pemetaan pilihan tindakan tetap. Untuk navigasi, ShapeNet, yang menyempurnakan bentuk angka ringkas dengan membandingkan untuk proses berpikir tentang ruang 3D, sangat mengurangi kesalahan perencanaan jalur. Karya-karya ini menunjukkan bahwa VLM dapat menggabungkan pengamatan dan proses berpikir dalam AI berwujud untuk menyelesaikan sejumlah besar tugas dengan pilihan tindakan tetap.

VLA untuk AI berwujud. VLA menggabungkan masukan dengan banyak jenis data dengan kendali tindakan tingkat rendah melalui alur kerja yang dapat dibedakan. Contoh karya seperti RT-2 pertama-tama mengkodekan citra robot saat ini, instruksi bahasa, dan tindakan robot pada langkah waktu tertentu dan mengubahnya menjadi potongan data teks; kemudian menggunakan LLM untuk proses berpikir berdasarkan makna dan pemecahan tugas; akhirnya, mende-tokenisasi potongan data yang dihasilkan menjadi tindakan akhir. Octo melakukan pelatihan awal pada 100.000 contoh tindakan robot dengan label penjelas bahasa, mencapai penggunaan alat lintas bentuk fisik. Untuk penggerakan benda yang cekatan, PerAct memanfaatkan gambaran kotak kecil 3D untuk mencapai ketepatan genggaman tingkat milimeter. Karya-karya ini menunjukkan bahwa VLA dapat bertindak pada seluruh sistem AI berwujud dan mencapai perluasan untuk menyesuaikan diri pada robot dan lingkungan baru.

Contoh Karya

MLLM dapat memberdayakan Pengamatan Aktif, Proses Berpikir Berwujud, dan Interaksi yang Menyesuaikan dari AI berwujud. MLLM untuk Pengamatan Aktif. Pertama, MLLM dapat meningkatkan 3D SLAM. Dengan menghubungkan pengamatan gambar ke dalam gambaran makna, MLLM melengkapi alur kerja SLAM tradisional dengan informasi sesuai

keadaan tingkat tinggi seperti kategori objek, relasi ruang, dan makna adegan. Karya representatif seperti SEO-SLAM memanfaatkan MLLM untuk menghasilkan label yang lebih spesifik dan deskriptif untuk objek, sambil terus memperbarui matriks kebingungan multikelas untuk mengurangi kecenderungan yang tidak seimbang dalam deteksi objek. Kedua, MLLM dapat meningkatkan Memahami Ruang 3D. Pengamatan berdasarkan kamera tetap menjadi pengaturan dominan dalam AI yang digerakkan oleh MLLM, karena masukan RGB selaras secara alami dengan pelatihan awal bahasa gambar dari banyak model dasar.

Karya representatif seperti EmbodiedGPT memanfaatkan sinergi ini untuk memetakan masukan gambar 2D ke dalam fitur yang kaya secara makna yang selaras dengan tujuan berdasarkan bahasa. Terakhir, MLLM dapat meningkatkan Penjelajahan Lingkungan Secara Aktif. MLLM juga telah merevolusi cara robot berinteraksi dengan lingkungannya, khususnya dalam interaksi siklus tertutup yang digerakkan oleh umpan balik. Karya representatif seperti LLM3 berfokus pada umpan balik tingkat gerak terstruktur, yang menggabungkan sinyal seperti deteksi tabrakan ke dalam loop perencanaan, memungkinkan model untuk secara berulang merevisi urutan tindakan simbolik. MART, di sisi lain, memanfaatkan umpan balik interaksi untuk meningkatkan kualitas pengambilan.

MLLM untuk Proses Berpikir Berwujud. Pertama, MLLM dapat meningkatkan perencanaan mandiri yang didorong oleh tugas. Sistem AI berwujud dengan MLLM dapat secara langsung memetakan tujuan tingkat tinggi ke urutan tindakan terstruktur, atau mengadopsi strategi perencanaan perantara yang terus berinteraksi dengan lingkungan untuk menyempurnakan rencana mereka. Karya representatif seperti CoT-VLA memperkirakan citra tujuan kecil perantara yang menggambarkan hasil yang diinginkan dari tugas kecil, membantu sistem AI memvisualisasikan dan menalar melalui setiap langkah dari tugas yang rumit. Kedua, MLLM dapat meningkatkan refleksi diri yang didorong oleh memori. MLLM memungkinkan sistem AI untuk belajar dari pengalaman menggunakan bagian sistem memori bawaan ini.

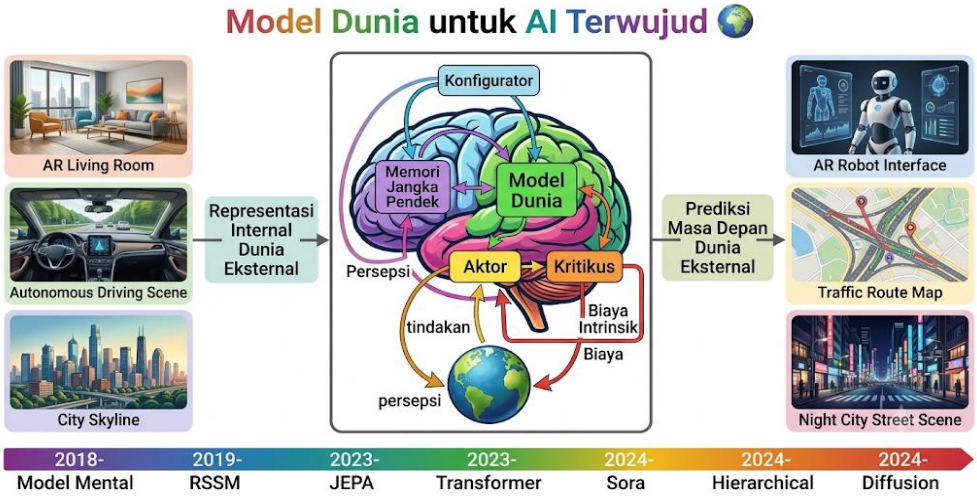
Karya representatif seperti Reflexion meningkatkan kinerja sistem AI melalui umpan balik linguistik yang dihasilkan sendiri, yang disimpan dalam tempat penyimpanan sementara berdasarkan kejadian dan dimanfaatkan untuk memandu perencanaan masa depan. Terakhir, MLLM dapat meningkatkan model dasar dengan banyak jenis data berwujud. MLLM dapat diadaptasi ke dunia fisik melalui pelatihan awal berkelanjutan atau penyempurnaan dalam pengaturan berwujud. Contoh karya yang representatif antara lain Qwen-VL dan InternVL, serta model yang mendukung keselarasan jenis data yang lebih luas, seperti Qwen2.5-Omni. MLLM untuk Interaksi yang Menyesuaikan. Pertama, MLLM dapat meningkatkan kendali tindakan. MLLM memiliki kemampuan untuk memecah tugas-tugas rumit menjadi tugas kecil yang dapat ditindaklanjuti. Untuk lebih lanjut menghasilkan sinyal kendali berkelanjutan untuk setiap tugas kecil, MLLM menghasilkan tindakan secara yang membuat keluaran secara berurutan secara berurutan atau menggunakan kepala strategi tambahan untuk memproses lebih lanjut gambaran internalnya. Kemajuan terbaru juga mengeksplorasi pembuatan kode yang dapat dilaksanakan dengan MLLM, memungkinkan robot untuk mengikuti strategi kendali yang dapat diinterpretasikan dan disesuaikan. Kedua, MLLM dapat meningkatkan interaksi perilaku. Melalui interaksi dengan lingkungan, MLLM juga mampu menghasilkan

urutan tindakan perilaku dalam satu langkah. Karya representatif seperti π-0 menggabungkan tulang punggung bahasa-penglihatan dengan dekoder pencocokan aliran untuk menghasilkan lintasan perilaku yang halus dan diperpanjang secara berdasarkan waktu.

Terakhir, MLLM dapat meningkatkan pengambilan keputusan bekerja sama. Salah satu jalur riset berfokus pada sistem banyak sistem AI yang bertujuan untuk mencapai koordinasi tingkat manusia dan menyesuaikan diri dengan cepat terhadap tantangan yang tidak terduga. Misalnya, Combo memperkenalkan kerangka kerja baru yang meningkatkan kerja sama antar sistem AI dibagi ke tiap bagian yang beroperasi hanya dengan pengamatan gambar egosentris. Upaya lain menyelidiki kerja sama manusia-sistem AI. VLAS mencontohkan hal ini dengan menyelaraskan perintah verbal manusia dengan konteks gambar melalui pengubah masukan ucapan dan MLLM gaya LLaVA, memungkinkan interaksi manusia-sistem AI yang lancar dan percakapan.

Model Dunia Meningkatkan AI berwujud

Model dunia memberdayakan AI berwujud dengan membangun gambaran di dalam sistem dan perkiraan keadaan mendatang dari dunia eksternal (seperti yang ditunjukkan pada Gambar 1.5), memudahkan interaksi berwujud yang sesuai dengan hukum fisika di lingkungan berubah-ubah. Gambaran di dalam sistem dari Dunia Eksternal. Gambaran di dalam sistem membantu mengompres data dari sensor mentah ke dalam ruang data tersembunyi yang terorganisir. Ruang-ruang ini menyimpan gerakan objek, hukum fisika, dan tata letak ruang, yang membantu sistem AI untuk mengetahui apa yang ada dan bagaimana perilakunya. Bentuk angka ringkas semacam itu menjaga lapisan hubungan yang jelas antara objek dan lingkungan sekitarnya, sesuai dengan bagaimana struktur dunia nyata terbentuk. Desain ini memudahkan penggunaan ide yang sama, seperti gravitasi atau permanensi objek, dalam pengaturan yang berbeda.



Gambar 1.5 Peta jalan pengembangan model dunia untuk AI berwujud

Mereka juga memungkinkan sistem AI untuk mengajukan pertanyaan "bagaimana jika" dengan menyimpan variabel terpisah untuk apa suatu objek dan bagaimana objek tersebut terhubung dengan hal-hal lain. Pemisahan ini memungkinkan bagian-bagian disesuaikan

secara luwes dalam pikiran. Hal ini juga meningkatkan pembelajaran, karena sistem AI dapat menggunakan kembali fitur bersama saat beralih tugas. Model dunia dengan ruang internal yang baik dapat memeriksa ketidakpastiannya sendiri dan mencari informasi yang hilang bila diperlukan. Dengan menyimpan aliran waktu dan struktur ruang, mereka menjaga rencana tetap realistis dan menyaring tindakan yang melanggar aturan fisik sebelum mencobanya. Pada akhirnya, ruang seperti itu memberikan dasar untuk pemikiran sebab-akibat, mirip dengan bagaimana manusia membangun model sederhana dari apa yang mereka rasakan.

Perkiraan keadaan mendatang Dunia Eksternal. Perkiraan keadaan mendatang memungkinkan sistem AI untuk memperkirakan kemungkinan imbalan untuk rangkaian tindakan dalam rentang waktu yang berbeda, sambil mengikuti aturan fisik. Pandangan ke depan ini menghubungkan langkah-langkah langsung dengan tujuan jangka panjang. Jalur yang tidak masuk akal, seperti berjalan menembus dinding, atau pilihan yang boros, seperti menggunakan sumber daya terlalu cepat, dapat dihindari sejak dini. Perkiraan jangka panjang membantu sistem AI menyesuaikan diri antara mencoba opsi baru dan tetap berpegang pada rencana yang aman. Dengan melihat ke depan, mereka dapat melewati pilihan lokal yang buruk tetapi tetap bertindak untuk mencapai tujuan jangka pendek. Ramalan ini juga mengukur ketidakpastian.

Ramalan ini membedakan pola yang stabil, seperti rutinitas harian, dari perubahan mendadak. Hal ini membuat perencanaan dalam kondisi berisiko menjadi lebih andal. Perkiraan simulasi menghemat sampel karena latihan mental membutuhkan biaya lebih sedikit daripada coba-coba di dunia nyata. Ini sangat penting untuk tugas-tugas di mana kesalahan sangat merugikan, seperti mobil mandiri atau robot medis. Pemeriksaan terus-menerus terhadap kesalahan perkiraan mendorong model untuk memperbaiki dirinya sendiri dari waktu ke waktu. Koreksi diri ini menjaga fisika internal mereka lebih dekat dengan kenyataan. Pada akhirnya, kemampuan untuk melihat ke depan ini memberi sistem AI buatan semacam pandangan ke depan yang praktis. Mereka beralih dari sekadar bereaksi menjadi bertindak dengan tujuan yang jelas.

Karya Representatif

AI berwujud dengan model dunia terutama dapat dibagi menjadi tiga struktur kritis: model dunia berdasarkan Recurrent State Space Model (RSSM) untuk AI berwujud, model dunia berdasarkan Joint-Bentuk angka ringkas Predictive Architecture (JEPA) untuk AI berwujud, dan model dunia berdasarkan Transformer untuk AI berwujud. Model dunia berdasarkan hierarki dan model dunia berdasarkan difusi mirip dengan struktur lainnya dan ditunjukkan pada Gambar 1.5. Model Dunia Berdasarkan RSSM untuk AI yang Diwujudkan. RSSM merupakan susunan sistem dasar yang mendasari keluarga langkah kerja Dreamer.

Kerangka kerja ini meningkatkan kemampuan perkiraan dalam gambaran data tersembunyi dengan memperoleh perubahan lingkungan berdasarkan waktu melalui masukan gambar, yang kemudian memungkinkan pemilihan tindakan melalui perbaikan lintasan tersembunyi. Melalui pemecahan ortogonal dari kondisi tersembunyi menjadi bagian berdasarkan peluang dan pasti, susunan sistem ini secara jelas memperhitungkan pola sistematis dan ketidakpastian lingkungan. Efektivitasnya yang telah terbukti dalam aplikasi

kendali gerak robot telah menginspirasi banyak studi turunan yang dibangun di atas kerangka teoritisnya.

Model Dunia berdasarkan JEPA untuk AI berwujud. JEPA menyediakan struktur untuk mengembangkan sistem kecerdasan mesin mandiri. Susunan sistem ini menetapkan hubungan pemetaan antara data masukan dan hasil yang diantisipasi melalui pembelajaran gambaran. Berbeda dari pendekatan yang menghasilkan data biasa, JEPA beroperasi di ruang data tersembunyi abstrak daripada menghasilkan pembentukan ulang piksel demi piksel, sehingga memprioritaskan pengambilan ciri makna daripada sintesis sinyal tingkat rendah. Dasar terkait cara kerja utama JEPA melibatkan cara pelatihan mandiri di mana jaringan saraf belajar untuk menyimpulkan segmen data yang terhalang atau tidak teramati. Pelatihan awal semacam itu pada kumpulan data tak berlabel yang ekstensif memungkinkan pembelajaran transfer di seluruh aplikasi hilir, menunjukkan peningkatan kemampuan untuk berlaku pada keadaan lain untuk bidang gambar dan non-gambar.

Model Dunia Berdasarkan Transformer untuk AI berwujud. Berasal dari riset pemrosesan bahasa alami, struktur Transformer pada dasarnya bergantung pada cara memilih bagian penting untuk memproses urutan masukan melalui pembobotan konteks yang diparalelkan. Desain ini memungkinkan perhitungan komputer bersamaan dari ketergantungan antar elemen, mengatasi batasan pemrosesan berurutan yang melekat pada Jaringan Saraf Berulang (RNN). Bukti empiris menunjukkan kinerja superior dalam bidang yang membutuhkan daya simpan memori persisten dan pengalamatan memori jelas untuk proses berpikir terkait proses berpikir, yang telah mendorong adopsinya dalam riset Belajar melalui imbalan sejak tahun 2020.

Kemajuan yang ada telah berhasil mengimplementasikan model dunia menggunakan varian Transformer, mengungguli susunan sistem RSSM dalam keadaan interaktif yang intensif memori. Terutama, kerangka kerja Genie Google menggunakan Transformer Ruang-Berdasarkan waktu (ST-Transformer) untuk menciptakan lingkungan interaktif buatan melalui pelatihan awal video mandiri skala besar. Kemajuan besar ini menetapkan cara baru untuk pembuatan model dunia yang dapat ditindaklanjuti, mengungkapkan potensi transformatif untuk lintasan pengembangan model dunia.

Hubungan Antara MLLM dan Model Dunia

MLLM memungkinkan proses berpikir tentang tugas sesuai keadaan tetapi mengabaikan batasan fisik, sementara model dunia unggul dalam simulasi yang mengikuti hukum fisika tetapi kurang memiliki pemahaman makna tingkat umum. Gabungan keduanya menjembatani kemampuan memahami makna dengan interaksi fisik yang mendasar. Keterbatasan MLLM untuk AI berwujud (tanpa model dunia). MLLM masih memiliki dua kekurangan utama dalam AI berwujud. Pertama, mereka sering kali gagal menyesuaikan perkiraan mereka pada fisika yang sebenarnya. Hal ini dapat menyebabkan rencana yang masuk akal dalam teks tetapi gagal di dunia nyata. Misalnya, mengabaikan gesekan atau tekstur permukaan dapat menyebabkan robot menjatuhkan objek atau tergelincir. Kedua, MLLM bereaksi buruk ketika lingkungan berubah dengan cepat. Mereka menangani pemecahan tugas dengan baik dalam bahasa tetapi kesulitan untuk mengubah tindakan secara

langsung. Alasan utamanya adalah mereka bergantung pada data pelatihan tetap. Mereka kekurangan umpan balik konstan yang berasal dari interaksi fisik nyata.

Keterbatasan Model Dunia untuk AI berwujud (tanpa LLM/MLLM). Model dunia juga memiliki keterbatasannya sendiri. Mereka lemah dalam menangani proses berpikir abstrak atau terbuka. Kekuatan utama mereka adalah menirukan dunia fisik, bukan memahami konteks yang lebih luas. Mereka juga kurang efektif ketika tugas membutuhkan langkah-langkah yang luwes. Tanpa pengetahuan bawaan, sulit bagi mereka untuk memecah tujuan besar menjadi bagian-bagian yang dapat dikerjakan. Misalnya, model dunia yang hanya mengenal bentuk kaku mungkin gagal jika harus berurusan dengan objek lunak atau luwes kecuali dilatih ulang dari awal. Perkiraan mereka sangat bergantung pada data dari tugas-tugas spesifik. Hal ini menyulitkan untuk menskalakannya ke bidang baru tanpa pelatihan lebih lanjut.

MLLM Meningkatkan Proses berpikir Model Dunia. Dengan memanfaatkan penyesuaian agar selaras antarjenis data dan kaitan dengan makna, MLLM memungkinkan model dunia untuk mengolah lingkungan rumit sesuai perubahan, meningkatkan proses berpikir berdasarkan makna, pemecahan tugas, dan interaksi manusia-robot. (1) MLLM dapat memperkaya model dunia dengan menggabungkan data gambar, suara, dan tekstual ke dalam gambaran makna yang terpadu. Misalnya, susunan sistem berbasis CLIP memungkinkan sistem AI untuk menyelaraskan adegan gambar dengan isyarat linguistik, mengurangi ketidakjelasan dalam pengenalan objek. (2) MLLM dapat meningkatkan kemampuan pemecahan tugas model dunia dengan mendekomposisi tujuan tingkat tinggi menjadi tugas kecil yang dapat dilaksanakan. Model seperti GPT-4V menghasilkan rencana langkah demi langkah menggunakan konteks lingkungan yang tersimpan dalam model dunia.

Untuk penggerakan benda oleh robot, Code-as-Policies menerjemahkan instruksi bahasa alami ke dalam cuplikan kode, memanfaatkan model dunia untuk melacak keadaan perantara. (3) MLLM memungkinkan model dunia untuk memperbaiki gambaran di dalam sistem melalui umpan balik manusia. Cara seperti Reinforcement Learning with Human Feedback (RLHF) memungkinkan sistem AI untuk memperbarui prioritas model dunia berdasarkan masukan korektif. Karya-karya dalam subbagian ini adalah semua cara yang mungkin bagi MLLM untuk meningkatkan proses berpikir model dunia, yang belum dicapai dalam karya-karya yang ada. Model Dunia Meningkatkan Interaksi MLLM. Model dunia dapat memainkan peran penting dalam memperbaiki MLLM dengan menyediakan hukum fisika, hubungan ruang-berdasarkan waktu, dan pengalaman interaksi siklus tertutup. Model dunia dapat mengurangi keterbatasan bawaan MLLM dalam kesinambungan dari waktu ke waktu dan dasar lingkungan, memungkinkan pengambilan keputusan yang lebih kuat dalam tugas-tugas berubah-ubah berwujud. (1) Model dunia dapat memberikan MLLM gambaran jelas hukum fisika (misalnya, gravitasi, gesekan) dan aturan akal sehat untuk membatasi usulan tindakan. Misalnya, Physion++, yang menggabungkan model biomekanik yang disimpan dalam model dunia, dapat digunakan untuk menyaring gerakan robot yang dihasilkan MLLM yang melanggar batas torsi; RoboGuide menyuntikkan peta keterisian ruang ke dalam perencana MLLM, mencegah tabrakan selama navigasi. (2) Model dunia dapat menstabilkan proses

berpikir MLLM dengan mempertahankan konteks ruang-berdasarkan waktu selama pemrosesan dengan banyak jenis data.

Misalnya, MemPrompt dapat menggunakan tempat penyimpanan sementara model dunia untuk menyelaraskan lintasan objek gambar dengan deskripsi linguistik, menyelesaikan ketidakjelasan di lingkungan yang berantakan; RoboMem dapat memanfaatkan perhatian yang diprioritaskan model dunia untuk menyaring kebisingan dari sensor yang tidak relevan, meningkatkan pemahaman adegan berdasarkan MLLM. (3) Model dunia dapat memungkinkan penyempurnaan berulang keluaran MLLM melalui interaksi siklus tertutup. Refleksi dapat menyimpan riwayat pelaksanaan tugas dalam model dunia, memungkinkan MLLM untuk mengoreksi kesalahan terkait gerakan menggunakan pola kegagalan. Karya-karya dalam subbagian ini merupakan berbagai cara yang memungkinkan bagi model-model dunia untuk meningkatkan pengambilan keputusan MLLM, yang belum tercapai dalam karya-karya yang ada sebelumnya.

Susunan sistem gabungan MLLM dan Model Dunia untuk AI berwujud

Kami mengusulkan susunan sistem gabungan MLLM dan model dunia untuk AI berwujud (seperti yang ditunjukkan pada Gambar 1.6), yang menjelaskan pentingnya mendalamnya dalam memungkinkan tugas-tugas rumit di dalam dunia fisik. Alur kerja spesifiknya adalah sebagai berikut, dengan panah yang menekankan proses pertukaran data. Robot Memasukkan Keadaan Diri MLLM/model dunia Perangkat Keras Bentuk fisik Robot. Proses dimulai dengan memasukkan keadaan diri yang melacak ukuran penilaian proprioseptif, seperti jumlah arah gerak, jumlah sensor, dll. Ukuran penilaian ini masuk ke dalam model dunia dan MLLM: model dunia menggunakannya untuk membangun gambaran di dalam sistem dari keadaan fisik sistem AI, sementara MLLM mengkontekstualisasikan keadaan ini untuk penyesuaian agar selaras tugas. Bentuk Perangkat Keras difokuskan pada penerapan model dunia dan MLLM ke dalam perangkat fisik untuk memecahkan masalah simulasi-ke-nyata. Alur dua arah ini memastikan tindakan menghormati batasan mekanis dan tujuan tingkat tinggi.

Gambar 1.6 AI berwujud yang diaktifkan oleh MLLM dan model dunia

MLLM Perencanaan Tugas model dunia Pembaruan Memori MLLM. MLLM memecah instruksi abstrak menjadi tugas kecil. Panah ke depan menyampaikan rencana ini ke model dunia, yang memperkirakan hasil berdasarkan pembuatan model lingkungan yang ada. Selama pelaksanaan, model dunia mencatat hasil ke dalam memori. Panah vertikal mengirimkan catatan ini ke bagian sistem pembaruan memori, yang menyusun memori menjadi pengalaman, menggambarkan pelupaan memori tugas masa lalu, pembaruan memori tugas saat ini, dan perkiraan memori tugas masa depan. Ini kemudian diumpankan kembali ke MLLM melalui panah, memperkaya basis pengetahuan mereka. Ini memungkinkan pembelajaran seumur hidup, di mana kegagalan masa lalu secara langsung menginformasikan perencanaan masa depan.

Lingkungan Pengamatan Aktif MLLM/model dunia Interaksi yang Menyesuaikan Lingkungan. Model dunia pertama-tama mendorong Pengamatan Aktif dengan memperkirakan perubahan lingkungan utama. Masukan dengan banyak jenis data kemudian digunakan untuk membangun gambaran di dalam sistem dunia eksternal melalui model dunia dan proses berpikir berdasarkan makna melalui MLLM. Kemudian, pemecahan tugas MLLM dan perkiraan keadaan mendatang model dunia memungkinkan pemilihan tindakan dan interaksi lingkungan. Pengamatan dapat menyesuaikan diri dan interaksi lingkungan berubah-ubah dicapai melalui pengulangan berkelanjutan.

Pengaturan model dunia-MLLM gabungan menawarkan kerangka kerja yang kuat untuk AI berwujud. Seperti yang ditunjukkan pada Tabel 1.4, MLLM menangani proses berpikir berdasarkan makna dengan baik. Mereka memecah tugas, mengikuti konteks, dan menyesuaikan rencana dengan menggunakan berbagai jenis masukan. Secara paralel, model dunia menambatkan rencana ini dalam realitas fisik. Mereka menirukan aturan dunia nyata sehingga tindakan tetap layak. Menggabungkan keduanya memungkinkan sistem AI untuk menghubungkan pemikiran abstrak dengan tindakan yang mematuhi fisika. Keseimbangan ini meningkatkan pilihan dalam kondisi yang berubah. Misalnya, MLLM dapat menyusun rencana. Model dunia kemudian memeriksa apakah rencana tersebut berfungsi dalam praktik. Siklus ini menyempurnakan rencana langkah demi langkah. Desain gabungan seperti itu juga membantu kemampuan berlaku pada keadaan lain di berbagai jenis data. Ini membuat sistem lebih kuat ketika beberapa informasi hilang atau situasinya baru. Hal ini dilakukan dengan menghubungkan simbol dengan sinyal sensor.

Tabel 1.4 Perbandingan kualitatif sistem MLLM-saja, model dunia-saja, & gabungan keduanya dalam AI terwujud.

Kinerja	Hanya LLM/MLLM	Hanya model dunia	Gabungan MLLM-model dunia
Pemahaman Makna	Keunggulan dalam proses berpikir tentang tugas sesuai keadaan dan pemahaman bahasa alami	Terbatas dalam pemahaman makna terbuka	Menggabungkan abstraksi pemahaman makna tingkat umum dengan penyesuaian agar selaras sesuai keadaan yang berdasar

Pemecahan tugas	Logika berurutan memungkinkan perencanaan tugas kecil melalui perintah bahasa	Kurangnya cara kerja pemecahan tugas yang dapat digeneralisasi	Rencana makna disempurnakan melalui kelayakan fisik via putaran perencanaan-pelaksanaan gabungan
Kepatuhan Fisika	Mengabaikan batasan dan perubahan fisik dalam interaksi dunia nyata	Simulasi sadar fisika dengan konsistensi berdasarkan waktu	Menegakkan penyesuaian agar selaras makna-fisik untuk rencana yang aman dan dapat dilaksanakan
Perkiraan keadaan mendatang	Kurangnya proses berpikir berdasarkan imajinasi	Perkiraan beberapa langkah cakrawala panjang dengan pembuatan model ketidakpastian	Menggabungkan pandangan jauh ke depan simbolis dan imajinasi berdasarkan fisik
Interaksi Langsung	Daya tanggap yang buruk terhadap umpan balik lingkungan dan waktu tunda proses berpikir yang besar	Mendukung kendali yang langsung memperkirakan keadaan berikutnya melalui simulasi keadaan masa depan	Memungkinkan penyesuaian *online* melalui penyempurnaan rencana berulang dan pembaruan memori
Struktur Memori	Memori yang jarang dan tidak terstruktur	Ruang data tersembunyi yang teratur menyandikan perubahan objek dan hubungan sebab-akibat	Menggabungkan memori makna dan pembuatan model dunia untuk pembelajaran sepanjang hayat dan refleksi
Kemampuan berkembang	Terbatas pada ruang tugas yang telah dilatih sebelumnya	Transfer yang buruk ke tugas yang belum pernah dilihat tanpa pelatihan ulang	Kemampuan berlaku pada keadaan lain lintas-tugas, lintas-bidang melalui sinergi simbolis dan sensorimotorik

Namun, pengaturan gabungan ini membawa tantangan baru. Pertama, MLLM (Model Linier Multidimensi) seringkali berjalan lambat saat melakukan proses berpikir, sementara model dunia harus melacak fisika secara real-time. Menjaga sinkronisasi keduanya sulit dan dapat memperlambat reaksi. Kedua, rencana yang dibuat oleh MLLM mungkin melanggar batasan fisik yang belum dibuat sebagai model oleh model dunia. Ketiga, model dunia terus memperbarui memori. Jika tidak ditangani dengan baik, hal ini dapat membebani MLLM dengan terlalu banyak detail. Selain itu, pelatihan sistem ini membutuhkan kumpulan data campuran yang sangat besar. Ini harus mencakup kasus-kasus ekstrem yang jarang terjadi untuk memastikan keamanan. Masih sulit untuk membuatnya tangguh terhadap sensor yang bising atau data yang hilang. Mengatasi masalah ini membutuhkan MLLM yang lebih cepat, umpan balik yang lebih erat antar bagian, dan cara yang lebih cerdas untuk memfilter konteks secara langsung.

1.5 APLIKASI UTAMA AI BERWUJUD

Kemajuan teknologi AI berwujud telah mendorong penerapannya di berbagai bidang aplikasi di mana interaksi fisik sangat penting. Dari robot mandiri yang menavigasi lingkungan yang tidak terstruktur hingga drone yang beroperasi di lingkungan udara yang berubah-ubah,

AI berwujud menunjukkan kemampuannya untuk menangani pengamatan, proses berpikir, dan kendali dalam konteks dunia nyata. Bagian ini menekankan aplikasi representatif, dengan fokus pada bagaimana AI berwujud memungkinkan sistem AI untuk mencapai otonomi yang kuat, kemampuan berlaku pada keadaan lain tugas, dan pelaksanaan yang aman. Dengan menganalisis kasus-kasus ini, kami menunjukkan bagaimana kecerdasan berwujud diterjemahkan dari model berdasarkan gagasan ke dalam sistem praktis dengan dampak sosial dan industri yang nyata.

Robot

Sistem AI berwujud secara aktif berinteraksi dengan dunia fisik dan mencakup berbagai bentuk berwujud, termasuk robot, peralatan pintar, kacamata realitas tertambah, dan kendaraan mandiri. Di antara semua ini, robot menonjol sebagai salah satu bentuk fisik yang paling menonjol dan banyak dipelajari. Tergantung pada keadaan aplikasi spesifik, robot dirancang dalam berbagai bentuk untuk memanfaatkan kemampuan perangkat kerasnya untuk tugas-tugas yang ditargetkan. Robot berwujud umumnya dapat dikategorikan ke dalam jenis-jenis berikut: (1) Robot dengan dasar tetap, seperti lengan robot, umumnya digunakan dalam otomatisasi laboratorium, pendidikan, dan manufaktur industri. (2) Robot beroda dikenal karena mobilitasnya yang tinggi dan banyak digunakan dalam logistik, pergudangan, dan inspeksi keamanan. (3) Robot beroda rantai menawarkan kemampuan off-road dan kemampuan gerak yang sangat baik, menunjukkan potensi yang kuat dalam pertanian, konstruksi, dan penanggulangan bencana. (4) Robot berkaki empat dihargai karena stabilitas dan kemampuan adaptasinya, menjadikannya ideal untuk eksplorasi di medan yang rumit, operasi penyelamatan, dan aplikasi militer. (5) Robot humanoid, yang dilengkapi dengan tangan yang cekatan, semakin banyak diterapkan dalam industri jasa, perawatan kesehatan, dan lingkungan bekerja sama manusia-robot. (6) Robot yang terinspirasi dari alam meniru gerakan dan fungsi organisme alami yang efisien, memungkinkan mereka untuk beroperasi secara efektif di lingkungan yang rumit dan berubah-ubah.

Robot Dengan dasar tetap. Robot dengan dasar tetap banyak digunakan dalam otomatisasi laboratorium, pelatihan pendidikan, dan manufaktur industri karena ketepatan tinggi, struktur kompak, dan stabilitas operasionalnya. Robot ini ditambatkan ke suatu basis, sehingga memungkinkan mereka mencapai ketepatan tingkat mikron untuk tugas-tugas yang membutuhkan pengulangan yang tepat. Dilengkapi dengan penggerak dan sensor yang ketepatan, robot dengan dasar tetap menawarkan kendali yang sangat baik, sehingga cocok untuk operasi yang rumit seperti sintesis kimia otomatis, perakitan sirkuit, dan bantuan bedah. Mereka juga sangat mudah diprogram, memungkinkan penyesuaian untuk berbagai aplikasi. Contoh yang terkenal termasuk Franka Emika Panda, KUKA iiwa, dan Sawyer oleh Rethink Robotics. Namun, ketidakmampuan robot dengan dasar tetap untuk bergerak membatasi jangkauan kerja dan kemampuan adaptasinya. Ketidakmampuan mereka untuk berpindah secara mandiri membatasi mereka pada lingkungan tetap dan membutuhkan sistem tambahan untuk tugas multi-lokasi. Keterbatasan ini juga mengurangi keluwesan bekerja sama mereka ketika diintegrasikan ke dalam tim manusia-robot yang berubah-ubah. Meskipun ketepatan dan keandalannya tinggi, robot-robot ini paling cocok untuk lingkungan di mana jangkauan

ruang tidak kritis, menekankan otomatisasi khusus tugas daripada keluwesan ruang. Robot Beroda. Robot beroda banyak diadopsi dalam logistik, pergudangan, dan keamanan karena mobilitasnya yang efisien dan kesederhanaan strukturnya. Dengan efisiensi energi yang tinggi, biaya rendah, dan navigasi cepat di permukaan datar, robot ini ideal untuk lingkungan terstruktur. Robot-robot ini biasanya menggabungkan sensor canggih seperti LiDAR dan kamera untuk memungkinkan perencanaan jalur mandiri dan penghindaran rintangan. Contoh yang menonjol termasuk robot Kiva yang digunakan di gudang Amazon dan robot Jackal oleh Clearpath Robotics. Kemampuan mereka untuk dengan cepat mengangkut barang dan berpatroli di ruang yang ditentukan telah merevolusi manajemen inventaris dan inspeksi keamanan.

Namun, kegunaannya terbatas di medan yang rumit karena ketergantungannya pada permukaan yang halus untuk kinerja optimal. Robot beroda memiliki kemampuan beban yang terbatas dan kesulitan dengan tugas-tugas yang membutuhkan orientasi ulang berubah-ubah atau gerak di tanah yang tidak rata. Meskipun mampu melakukan kerja sama terbatas di dalam ruangan dengan manusia, kemampuan penyesuaian mereka di lingkungan yang tidak terstruktur atau di luar ruangan jauh lebih rendah daripada platform berkaki atau beroda rantai. Akibatnya, robot beroda paling cocok untuk aplikasi berulang dan berkapasitas tinggi di mana medan yang tetap dan infrastruktur navigasi yang andal tersedia. Robot Beroda Rantai. Robot beroda rantai menawarkan kemampuan off-road yang unggul dan sangat cocok untuk lingkungan yang menantang seperti pertanian, konstruksi, dan keadaan pemulihan bencana. Desain rantai berkelanjutan mereka meningkatkan kontak dengan tanah, mengurangi tekanan dan meningkatkan stabilitas di medan yang lunak atau tidak rata. Hal ini memungkinkan perlintasan yang efektif melalui lumpur, pasir, puing-puing, dan tanjakan di mana robot beroda gagal. Sistem beroda rantai sering menggabungkan suspensi dan cara kerja penggerak yang kuat, memungkinkan mereka untuk mempertahankan traksi sambil menopang muatan berat. Kemampuan ini membuat mereka cocok untuk operasi berisiko tinggi, termasuk pengintaian militer dan penjinakan bahan peledak, seperti yang dicontohkan oleh PackBot dari iRobot.

Terlepas dari keunggulan ini, robot beroda rantai cenderung mengalami efisiensi energi yang rendah karena gesekan permukaan yang tinggi. Kecepatan maksimum mereka di permukaan datar biasanya lebih rendah daripada robot beroda, dan kemampuan gerak mereka di ruang terbatas terbatas. Selain itu, sistem rantai yang rumit membutuhkan banyak tenaga mekanis, sehingga meningkatkan biaya perawatan. Oleh karena itu, meskipun robot beroda rantai unggul dalam kondisi lapangan yang menuntut, kepraktisannya berkurang dalam keadaan yang memprioritaskan kecepatan, efisiensi, atau kemampuan menyesuaikan susunan bagian. Mereka tetap sangat diperlukan dalam peran yang memprioritaskan kekokohan dan kemandirian medan daripada ketepatan atau kelincahan.

Robot Berkaki Empat. Robot berkaki empat, yang dibuat sebagai model seperti hewan berkaki empat, memberikan mobilitas dan stabilitas yang sangat baik di medan yang tidak rata. Struktur anggota tubuhnya yang multi-sendi memungkinkan mereka untuk melintasi lingkungan yang tidak rata, memanjat rintangan, dan menjaga keseimbangan dalam kondisi berubah-ubah. Hal ini membuat mereka ideal untuk misi penyelamatan, eksplorasi, dan

aplikasi militer, terutama di daerah yang tidak dapat diakses oleh roda atau trek. Platform terkemuka termasuk A1 dan Go1 dari Unitree Robotics, Spot dari Boston Dynamics, dan ANYmal C dari ANYbotics. Robot-robot ini menggabungkan sistem sensor seperti LiDAR dan penglihatan dua kamera dengan Belajar melalui imbalan atau kendali berdasarkan model untuk mencapai perilaku dapat menyesuaikan diri. Robot Spot, misalnya, digunakan untuk inspeksi industri dan misi mandiri di lokasi berbahaya. Desain terbagi menjadi beberapa bagian dan navigasi mandiri ANYmal C cocok untuk operasi luar ruangan yang lama dan eksplorasi ruang angkasa. Namun, robot berkaki empat biasanya mahal karena desain mekanik dan sistem persepsinya yang rumit. Mereka juga menghadapi keterbatasan daya tahan baterai, seringkali membutuhkan pengisian ulang yang sering untuk penggunaan yang lama. Terlepas dari tantangan ini, kemampuan mereka untuk menavigasi lingkungan yang berantakan dan tidak terstruktur membuat mereka sangat berharga dalam keadaan yang menuntut keluwesan medan dan pengambilan keputusan semi-mandiri.

Robot Humanoid. Robot humanoid meniru bentuk dan perilaku manusia, memungkinkan mereka untuk melakukan tugas-tugas rumit di lingkungan bekerja sama, perawatan kesehatan, dan berfokus layanan. Desain antropomorfik mereka mencakup tangan yang cekatan dengan banyak jumlah arah gerak dan penggabungan sensor, memungkinkan penggerakan benda halus dan interaksi seperti manusia. Atlas oleh Boston Dynamics menunjukkan gerakan yang sangat berubah-ubah seperti berlari dan membalik, sementara ASIMO Honda dapat berjalan, menaiki tangga, dan berinteraksi secara sosial. Seri HRP dari AIST berfokus pada kerja sama manusia-robot, khususnya di lingkungan industri. Selain itu, Pepper dari Softbank Robotics menawarkan pengenalan emosi dan pemrosesan bahasa alami, yang banyak digunakan dalam layanan pelanggan dan pendidikan.

Meskipun serbaguna, robot humanoid menghadapi tantangan dalam menjaga stabilitas dan keandalan di lingkungan dunia nyata. Gerakan bipedal yang ketepatan dan kendali tangan yang halus membutuhkan langkah kerja kendali yang rumit dan loop umpan balik langsung. Selain itu, sistem hidrolik tradisional berukuran besar dan mahal untuk dipelihara. Upaya terbaru, seperti robot humanoid berdasarkan motor dari Tesla dan Unitree, bertujuan untuk meningkatkan kinerja dan kemampuan berkembang. Dengan penggabungan model bahasa yang besar, robot humanoid diharapkan dapat melakukan tugas yang semakin beragam secara mandiri, mengisi kesenjangan tenaga kerja di sektor-sektor yang membutuhkan.

Robot Biomimetik. Robot biomimetik bertujuan untuk meniru gerakan dan sifat fungsional organisme alami untuk melakukan tugas di lingkungan yang rumit dan berubah-ubah. Robot ini sering menggunakan material lunak dan luwes serta desain hemat energi untuk meniru pergerakan biologis. Aplikasinya mencakup pemantauan lingkungan, biomedis, dan riset ekologi. Robot yang terinspirasi ikan menunjukkan pergerakan bawah air yang senyap dan lincah, ideal untuk eksplorasi laut. Robot seperti serangga unggul dalam perilaku berkelompok dan dapat mengakses ruang terbatas untuk inspeksi struktural. Robot berbadan lunak dapat berinteraksi dengan aman dengan struktur yang halus, sehingga cocok untuk prosedur medis minimal invasif. Desain ini mengurangi kerumitan mekanis dan memungkinkan

pengoperasian di lingkungan yang terlalu sensitif untuk sistem kaku. Namun, robot biomimetik menghadapi keterbatasan yang besar. Proses desain dan fabrikasinya menuntut secara teknis dan mahal, sehingga menghambat produksi skala besar. Selain itu, penggunaan material lunak seringkali mengurangi daya tahan, terutama dalam kondisi yang keras atau ekstrem. Meskipun menjanjikan dalam aplikasi khusus, penggunaan yang lebih luas membutuhkan kemajuan dalam ilmu material, penyimpanan energi, dan ketepatan kendali. Terlepas dari tantangan ini, robot biomimetik menawarkan arah yang menarik bagi robotika yang memprioritaskan kemampuan menyesuaikan diri, keselamatan, dan harmoni lingkungan.

Drone

Dalam lingkungan yang berubah-ubah, kendaraan udara tak berawak (UAV) umumnya bergantung pada kamera dan berbagai sensor yang terpasang untuk menangkap citra dan aliran video secara real-time, yang mendukung pengambilan keputusan tingkat tinggi. Untuk melakukan tugas pemahaman makna, UAV harus melampaui sekadar platform pengumpulan data. Mereka semakin diharapkan untuk menunjukkan kecerdasan berwujud, memungkinkan mereka untuk membuat model lingkungan secara makna dan berinteraksi melalui bahasa alami. Hal ini memungkinkan UAV untuk menghasilkan respons makna yang bermakna yang selaras dengan perintah saat digunakan manusia berdasarkan masukan pengamatan mereka. Kemajuan terbaru dalam Model Dasar (FM), khususnya Model AI Bahasa Besar (LLM), telah menunjukkan otonomi dan kemampuan menyesuaikan diri yang kuat di berbagai bidang, termasuk kecerdasan berwujud. Sejumlah riset yang berkembang telah mengeksplorasi pengintegrasian LLM untuk meningkatkan pengamatan dan proses berpikir UAV dalam pengaturan berubah-ubah. Misalnya, de Zarzà dan kawan-kawan. dan de Curtò dan kawan-kawan. mengembangkan sistem pemahaman video UAV dengan menggabungkan beberapa model dasar. Metode-metode ini telah membuat kemajuan yang besar dalam pembuatan model makna.

Namun, mereka seringkali kekurangan kemampuan perencanaan tugas yang jelas, sehingga mengakibatkan kendali yang terbatas atas konsistensi respons. Selain itu, karena tidak adanya cara kerja untuk memanggil alat eksternal dan mengatur bagian terbagi menjadi beberapa bagian, sistem ini kesulitan untuk berlaku pada keadaan lain secara efektif dalam tugas-tugas rumit dan terbuka. Untuk mengatasi keterbatasan ini, cara kecerdasan buatan agenik yang muncul menawarkan kerangka kerja yang lebih luwes dan dapat digeneralisasi. Dibangun di sekitar LLM, cara ini mendukung fungsi-fungsi penting seperti pemecahan tugas, pemanfaatan alat, dan koordinasi antar beberapa sistem AI. Ini memberikan arah yang menjanjikan untuk memungkinkan UAV melakukan tugas pemahaman makna di lingkungan berubah-ubah secara terpadu dan mudah dikembangkan. AirVista-II mencontohkan pendekatan ini dengan membangun sistem AI untuk UAV yang berwujud.

Sistem ini mengorganisir adegan berubah-ubah ke dalam tiga kategori tugas: gambar, video pendek, dan video panjang, dan merancang alur kerja sistem AI dan strategi pemanggilan model yang disesuaikan untuk masing-masing kategori. Desain ini memungkinkan alur kerja mandiri lengkap yang mencakup pengenalan tugas, pemrosesan data, penafsiran makna, dan proses berpikir tingkat tinggi. Selain itu, riset dalam kecerdasan

buatan berwujud semakin mendapat perhatian. Salah satu tugas inti di bidang ini adalah Navigasi dengan Penglihatan dan Bahasa (VLN), yang berfokus pada memandu sistem AI ke lokasi target berdasarkan instruksi bahasa alami dan pengamatan gambar. Baru-baru ini, AerialVLN dan Open-UAV telah membuat kemajuan besar dengan memanfaatkan program simulasi UAV untuk mengisi kesenjangan dalam kumpulan data VLN udara. Upaya ini telah sangat memajukan bidang ini, tetapi masih menghadapi tantangan seperti keragaman data yang terbatas dan skala yang tidak memadai. Untuk mengatasi masalah ini, OpenFly telah diperkenalkan sebagai platform khusus untuk pengumpulan data skala besar dalam Navigasi dengan Penglihatan dan Bahasa udara. OpenFly menggabungkan beberapa mesin pembuatan gambar dan menyediakan rangkaian alat pembuatan data otomatis, memungkinkan pengumpulan data VLN udara yang beragam dan berkualitas tinggi secara efisien.

1.6 SIMULATOR, KUMPULAN DATA, DAN UKURAN PENILAIAN

Pengembangan dan penilaian sistem AI berwujud membutuhkan alat yang andal untuk eksperimen, pengujian dengan patokan, dan pemeriksaan. Karena penggunaan di dunia nyata seringkali mahal dan berisiko, program simulasi memainkan peran penting dengan menyediakan platform pengujian yang aman dan mudah dikembangkan. Sebagai pelengkap, kumpulan data yang dikurasi dan ukuran penilaian yang dirancang dengan baik memastikan kemampuan mengulang hasil dan memungkinkan perbandingan yang adil antar metode. Bagian ini meninjau program simulasi, kumpulan data, dan ukuran penilaian yang umum digunakan dalam riset AI berwujud, menekankan tingkat pentingnya dalam menjembatani inovasi berdasarkan langkah kerja dengan penilaian kinerja praktis. Bersama-sama, sumber daya ini membentuk dasar terkait cara kerja untuk memajukan bidang ini secara sistematis dan ketat.

Program simulasi untuk AI berwujud

Program simulasi berwujud memainkan peran sentral dalam kemajuan AI berwujud dengan menyediakan kerangka kerja yang aman, efisien, dan mudah dikembangkan untuk mengembangkan dan menilai sistem AI cerdas. Platform simulasi ini menghilangkan risiko dan biaya yang terkait dengan pengujian fisik, terutama dalam keadaan yang berbahaya atau sulit direproduksi di dunia nyata. Platform ini memungkinkan para ahli untuk bereksperimen dengan perilaku rumit secara terkontrol dan berulang, sehingga ideal untuk tugas-tugas seperti penggerakan benda oleh robot, navigasi mandiri, dan interaksi manusia-sistem AI. Selain keamanan dan keluwesan, program simulasi memudahkan pembuatan data skala besar untuk pelatihan dan penilaian. Program simulasi mendukung eksperimen dengan throughput tinggi, memungkinkan para ahli untuk mengeksplorasi berbagai kondisi lingkungan dan susunan sistem AI. Hal ini mempercepat pengembangan langkah kerja pembelajaran, mendukung studi pengujian dengan melepas bagian terperinci, dan memastikan kemampuan mengulang hasil melalui patokan uji standar dan lingkungan virtual bersama, yang penting untuk perbandingan yang bermakna di berbagai metode dan sistem.

Efektivitas program simulasi berwujud sangat bergantung pada kemampuannya untuk membuat model lingkungan yang realistis. Ini termasuk simulasi akurat dari sifat fisik, seperti

gravitasi, gesekan, dan perubahan objek, serta jenis data dari sensor seperti penglihatan, pendengaran, dan sentuhan. Tingkat realisme yang tinggi sangat penting untuk menjembatani kesenjangan antara simulasi dan penggunaan di dunia nyata, memastikan bahwa sistem AI yang dilatih dalam pengaturan virtual dapat berhasil mentransfer kemampuan mereka ke lingkungan fisik. Bagian ini memperkenalkan platform simulasi AI yang umum digunakan dalam dua kategori. Yang pertama terdiri dari program simulasi umum yang dibangun di atas program peniru hukum fisika yang memungkinkan kustomisasi tugas dan lingkungan yang ekstensif. Yang kedua mencakup program simulasi berdasarkan pembentukan ulang adegan dunia nyata, yang memprioritaskan fotorealisme dan ketepatan struktural untuk mendukung penilaian yang sangat mirip dengan kondisi dunia nyata.

Program simulasi Umum

Meskipun kekayaan interaksi fisik dan perubahan yang terus terjadi di lingkungan dunia nyata tidak tergantikan, penggunaan model berwujud di dunia fisik tetap mahal dan menghadirkan banyak tantangan praktis. Ini termasuk biaya perangkat keras, pemeliharaan, masalah keamanan, dan kemampuan berkembang yang terbatas. Akibatnya, program simulasi umum telah menjadi alat yang sangat diperlukan dalam AI berwujud. Program simulasi umum menawarkan beberapa manfaat inti. Mereka mendukung eksperimen yang dapat diulang, memungkinkan pembuatan prototipe yang cepat, dan menghasilkan data buatan untuk tugas pengamatan, kendali, dan pengambilan keputusan. Dengan menawarkan pengamatan dan kendali penuh atas nilai pengatur lingkungan, mereka memudahkan analisis terperinci dan pengujian langkah kerja dalam kondisi yang tetap.

Di antara platform yang banyak digunakan, Isaac Sim adalah program simulasi canggih yang dikembangkan oleh NVIDIA, dirancang untuk mendukung riset robotika dan AI dalam skala besar. Fitur-fiturnya meliputi simulasi fisik ketepatan tinggi, pelacakan sinar langsung untuk pembuatan gambar sangat mirip foto nyata, dan penggabungan tanpa batas dengan kerangka kerja pembelajaran mendalam. Isaac Sim mendukung berbagai aplikasi, termasuk simulasi kendaraan mandiri, robotika industri, dan kerja sama manusia-robot. Gazebo adalah program simulasi sumber terbuka terkemuka lainnya, terutama dihargai karena kompatibilitasnya dengan Robot Operating System (ROS). Ia menawarkan serangkaian alat yang kaya untuk membuat model perubahan robot, menirukan sensor seperti kamera dan lidar, dan menguji strategi koordinasi multi-robot. Gazebo umumnya digunakan dalam navigasi robot bergerak, pengembangan sistem kendali, dan riset akademis karena ekstensibilitasnya dan dukungan komunitas yang besar. PyBullet, antarmuka Python untuk program peniru hukum fisika Bullet, dikenal karena kesederhanaan dan aksesibilitasnya. Ia mendukung simulasi fisik langsung dengan perubahan benda kaku, deteksi tabrakan, dan batasan. PyBullet sangat cocok untuk tugas-tugas yang membutuhkan pengulangan cepat, seperti Belajar melalui imbalan dan penggerakan benda oleh robot. Ia juga tergabung dengan baik dengan kerangka kerja pembelajaran mendalam, memungkinkan simulasi strategi kendali berdasarkan pengamatan.

Tabel 1.5 Program simulasi umum (√: unggul). (HFPS: fisik presisi, HQGR: grafis tinggi, RRL: ragam robot, DLS: belajar mendalam, LSPC: komputasi paralel besar, ROS: integrasi ROS, MSS: multi-sensor, CP: lintas platform, Nav: navigasi, AD: kemudi otomatis, RL: belajar berimbalan, LSPS: simulasi paralel besar, MR: multi-robot, RS: simulasi robot, PE: peniru fisika, MLA: AI ML, OBSD: ODE, Bullet, Simbody, DART, BOVN: Bullet, ODE, Vortex, Newton, DAR: drone, AD, RL).

Program simulasi	Tahun	HFPS	HQGR	RRL	DLS	LSPC	ROS	MSS	CP	PE	Aplikasi
Isaac Sim	2023	√	√	√	√	√	√	√	√	PhysX	Nav, AD
Isaac Gym	2019		√		√	√				PhysX	RL, LSPS
Gazebo	2004		√	√			√	√	√	OBSD	Nav, MR
PyBullet	2017				√				√	Bullet	RL, RS
Webots	1996		√	√				√	√	ODE	RS
MuJoCo	2012	√			√				√	Kustom	RL, RS
Unity MLA	2017		√		√				√	Kustom	RL, RS
AirSim	2017		√						√	Kustom	DAR
MORSE	2015							√	√	Bullet	Nav, MR
V-REP	2013		√	√				√	√	BOVN	MR, RS

Tabel 1.5 merangkum fitur inti dan bidang aplikasi dari sepuluh program simulasi serbaguna yang sering digunakan dalam AI berwujud. Masing-masing menawarkan kekuatan yang berbeda dalam membuat model interaksi fisik, umpan balik dari sensor, atau penggabungan dengan kerangka kerja AI. Dengan memilih program simulasi yang paling tepat berdasarkan tujuan riset mereka, pengembang dapat mempercepat pembuatan prototipe, pengujian, dan penggunaan sistem kecerdasan berwujud di berbagai keadaan dunia nyata.

Program simulasi Berdasarkan Adegan Nyata

Untuk memajukan pengembangan sistem AI yang mampu melakukan tugas-tugas rumah tangga yang rumit, program simulasi ketepatan tinggi yang berdasarkan pada lingkungan dunia nyata menjadi sangat penting. Sistem-sistem AI ini harus memahami nuansa kehidupan sehari-hari manusia dan menjalankan tugas-tugas seperti navigasi, penggerakan benda, dan interaksi dalam pengaturan dalam ruangan yang rumit. Akibatnya, permintaan akan program simulasi yang menawarkan adegan fotorealistis, interaksi objek yang realistis, dan beragam jenis data dari sensor telah meningkat dengan jelas. Banyak dari program simulasi ini bergantung pada pemindaian 3D, fotogrametri, dan mesin game seperti Unity atau Unreal Engine untuk menciptakan kembali kondisi dunia nyata, menjadikannya platform ideal untuk melatih dan menilai sistem AI yang mampu melakukan tugas-tugas tersebut.

AI2-THOR, yang dikembangkan oleh Allen Institute for AI, adalah program simulasi berdasarkan Unity yang menawarkan lebih dari 2000 objek interaktif dan berbagai jenis adegan rumah tangga. Lingkungannya dibagi menjadi iTHOR, yang mendukung interaksi objek yang kaya, dan RoboTHOR, yang menyelaraskan tata letak virtual dengan ruang dunia nyata. SAPIEN menekankan interaksi fisik yang halus. Dibangun di atas mesin PhysX dan kumpulan

data PartNet-Mobility, ia memungkinkan sistem AI berwujud untuk memanipulasi sendi objek menggunakan kendali berdasarkan gaya yang realistis, yang penting untuk tugas penggerakan benda yang tepat. iGibson, yang dikembangkan oleh Stanford, berfokus pada beragam properti fisik, termasuk suhu, kelembaban, dan kebersihan. Ia menyediakan masukan dari sensor yang kaya (termasuk LiDAR) dan kendali motorik yang halus.

TDW dari MIT menggabungkan pembuatan gambar dengan banyak jenis data (gambar dan suara) dan mendukung interaksi dengan material kaku, lunak, dan cair, menawarkan keluwesan untuk beragam kebutuhan simulasi. Platform penting lainnya termasuk VirtualHome, yang memperkenalkan grafik lingkungan untuk perencanaan tingkat tinggi berdasarkan deskripsi bahasa; Habitat, dikenal karena kecepatannya, kemudahan untuk diperluas terbuka, dan dukungan kuat untuk navigasi; dan Matterport3D, sebuah kumpulan data skala besar yang memungkinkan navigasi berpandangan lebar berwujud. Perkembangan terbaru dalam konstruksi simulasi otomatis meliputi RoboGen, yang menggunakan LLM untuk menghasilkan adegan dari aset, HOLODECK, yang membangun adegan dari instruksi di AI2-THOR, PhyScene untuk pembangkitan adegan berdasarkan difusi, dan ProcTHOR untuk pembuatan adegan yang beragam dan mudah dikembangkan. Tabel 1.6 membandingkan program simulasi ini, menekankan fitur inti dan kasus penggunaan utamanya di seluruh navigasi, interaksi, dan perencanaan berwujud.

Tabel 1.6 Perbandingan program simulasi berdasar adegan nyata (Sensor S: Makna, L: LiDAR, A: Audio, CA: Tindakan berkelanjutan, 3D-SS: Pemindaian adegan 3D, MA: Beberapa sistem AI, OS: Status objek, 3D-A: Aset 3D)

Platform Simulasi	Tahun	Jumlah Adegan	CA	3D SS	Sensor	Fisika	MA	OS	3D-A
AI2-THOR	2017	120	✓	×	RGB-D, S	✓	✓	✓	✓
Matterport 3D	2018	90	×	✓	RGB-D, S	×	×	×	×
Habitat	2019	1000+	✓	✓	RGB-D, S	✓	✓	×	✓
Virtual Home	2018	50	×	×	RGB-D, S	✓	✓	✓	✓
SAPIEN	2020	46	✓	×	RGB-D, S	✓	×	✓	✓
iGibson	2021	15	✓	✓	RGB-D, S, L	✓	×	✓	✓
TDW	2021	-	✓	✓	RGB-D, S, A	✓	✓	✓	✓

Kumpulan data dan Ukuran penilaian untuk AI berwujud

Kumpulan data dan ukuran penilaian membentuk dasar empiris untuk riset AI berwujud. Sementara program simulasi menyediakan lingkungan terkontrol untuk pengujian, kumpulan data menyediakan aliran sensor dunia nyata atau buatan yang menjadi patokan uji pembelajaran dan kemampuan berlaku pada keadaan lain. Ukuran penilaian lebih lanjut memastikan perbandingan yang adil dengan mengukur aspek-aspek seperti ketepatan, daya tahan, efisiensi, dan kemampuan penyesuaian. Subbagian ini meninjau kumpulan data dan protokol penilaian yang banyak digunakan, menekankan pentingnya dalam memvalidasi sistem AI berwujud dan memandu kemajuan menuju praktik riset yang terstandarisasi dan dapat direproduksi.

Navigasi dengan Penglihatan dan Bahasa

Navigasi dengan Penglihatan dan Bahasa (VLN) menghadirkan beragam pengaturan tugas tergantung pada sifat instruksi bahasa dan kerumitan perilaku sistem AI. Instruksi dapat bervariasi dari arahan langkah demi langkah yang terperinci hingga tujuan tingkat tinggi atau bahkan tuntutan manusia yang samar. Demikian pula, kerumitan tugas berkisar dari satu tujuan navigasi hingga rencana beberapa langkah yang melibatkan interaksi, eksplorasi, dan tugas kecil berurutan. Perbedaan ini menghasilkan berbagai tantangan riset, yang mengarah pada pengembangan beberapa kumpulan data patokan uji yang dirancang untuk menilai sistem AI berwujud dalam kondisi yang berbeda, seperti yang ditunjukkan pada Tabel 1.7. Kumpulan data Room-to-Room (R2R), yang dibangun di atas Matterport3D, adalah patokan uji dasar di mana sistem AI mengikuti instruksi terperinci melalui grafik navigasi terpisah. Sistem AI harus menyelaraskan bahasa dan pengamatan gambar untuk memilih node berikutnya dan mencapai lokasi tujuan.

Tabel 1.7 Perbandingan kumpulan data VLN (I: Dalam Ruangan; O: Luar Ruangan; C: Berkelanjutan; D: Terpisah)

Kumpulan data	Tahun	Program simulasi	Lingkungan	Fitur	Ukuran
R2R	2018	Matterport3D	I, D	Instruksi langkah demi langkah	21,567
R4R	2019	Matterport3D	I, D	Instruksi langkah demi langkah	200,000+
VLN-CE	2020	Habitat	I, C	Instruksi langkah demi langkah	-
TOUCHDOWN	2019	-	O, D	Instruksi langkah demi langkah	9,326
REVERIE	2020	Matterport3D	I, D	Navigasi tujuan yang dideskripsikan	21,702
SOON	2021	Matterport3D	I, D	Navigasi tujuan yang dideskripsikan	3,848
DDN	2023	AI2-THOR	I, C	Navigasi berdasarkan permintaan	30,000+
ALFRED	2020	AI2-THOR	I, C	Navigasi dengan interaksi	25,743
OVMM	2023	Habitat	I, C	Navigasi dengan interaksi	7,892
BEHAVIOR-1K	2023	OmniGibson	I, C	Navigasi rentang panjang dengan interaksi	1,000
CVDN	2020	Matterport3D	I, D	Dialog, *oracle*	2,050
DialFRED	2022	AI2-THOR	I, C	Dialog, *oracle*	53,000

Room-for-Room (R4R) memperluas pengaturan ini dengan meningkatkan panjang jalur, menuntut memori jangka panjang dan dasar bahasa yang lebih efektif. VLN-CE memperkenalkan kendali berkelanjutan dalam lingkungan sangat mirip foto nyata, meningkatkan kesulitan pengambilan keputusan tindakan. Di luar adegan dalam ruangan,

TOUCHDOWN menempatkan sistem AI di lingkungan luar ruangan perkotaan menggunakan Google Street View. Sistem AI harus mengikuti instruksi rumit untuk menemukan objek target di lingkungan skala kota yang realistis, membutuhkan proses berpikir tentang ruang dalam skala besar. REVERIE juga menggunakan Matterport3D tetapi berfokus pada pengenalan tujuan tingkat tinggi. Sistem AI diharuskan untuk mencari dan berinteraksi dengan objek target yang tidak terlihat berdasarkan deskripsi bahasa alami yang ringkas. SOON menekankan mengikuti instruksi kasar-ke-halus, membimbing sistem AI untuk mempersempit area pencarian mereka secara progresif. DDN lebih jauh mengabstraksikan bentuk instruksi dengan mengganti target jelas dengan kebutuhan manusia umum, menantang sistem AI untuk menyimpulkan objek dan perilaku yang sesuai dari isyarat sesuai keadaan.

Dalam pengaturan berfokus tugas, ALFRED menggabungkan navigasi dan penggerakan benda menggunakan instruksi bertingkat dalam AI2-THOR. OVMM menilai tugas transportasi objek di berbagai adegan yang belum pernah dilihat sebelumnya, dengan simulasi dan penggunaan di dunia nyata. BEHAVIOR-1K mendorong batas dengan tugas-tugas jangka panjang yang kaya keterampilan yang berakar pada aktivitas manusia sehari-hari menggunakan platform OmniGibson. Navigasi berdasarkan dialog juga dieksplorasi dalam CVDN dan DialFRED, di mana sistem AI dapat mengajukan pertanyaan untuk menyelesaikan ketidakpastian atau meminta bantuan selama navigasi atau penggerakan benda. Kumpulan data ini memperkenalkan pembelajaran interaktif dan pengambilan keputusan, yang mendorong interaksi sistem AI-lingkungan-manusia yang lebih dalam.

Ukuran penilaian. Tergantung pada sifat tugas, ukuran penilaian yang berbeda menekankan pencapaian tujuan atau kepatuhan jalur. Secara umum, ukuran penilaian dapat dikategorikan menjadi ukuran yang berfokus pada tujuan, yang berfokus pada kemampuan sistem AI untuk mencapai target, dan ukuran penilaian kesetiaan jalur, yang menilai seberapa baik sistem AI mengikuti jalur yang diinginkan. Ukuran yang berfokus pada tujuan. Fokus utama pada penilaian seberapa efektif sistem AI mencapai atau mendekati tujuan yang ditentukan. Ukuran penilaian paling dasar adalah Tingkat Keberhasilan (SR), yang menghitung persentase episode di mana sistem AI berakhir dalam jarak ambang batas yang telah ditentukan dari target. Sementara SR memberikan ukuran biner penyelesaian tugas, Kemajuan Tujuan menawarkan penilaian berkelanjutan dengan mengukur pengurangan jarak ke tujuan selama suatu episode, menangkap keberhasilan parsial bahkan ketika tugas belum sepenuhnya selesai. Panjang Jalur (PL) mewakili total jarak yang ditempuh oleh sistem AI, berfungsi sebagai ukuran efisiensi.

Dalam praktiknya, jalur yang lebih pendek seringkali lebih disukai, terutama dalam aplikasi robotika dunia nyata di mana jalur yang lebih panjang menimbulkan biaya waktu dan energi yang lebih besar. Untuk menyeimbangkan penyelesaian tugas dengan efisiensi, Success weighted by Path Length (SPL) menyesuaikan tingkat keberhasilan berdasarkan rasio jalur terpendek terhadap jalur sebenarnya yang diambil. Demikian pula, Success weighted by Edit Distance (SED) membandingkan urutan tindakan sistem AI dengan urutan tindakan ahli, menggabungkan keberhasilan tugas dan ketepatan tingkat tindakan. Untuk memperhitungkan keadaan hampir berhasil, Oracle Navigation Error (ONE) mengukur jarak terpendek antara

tujuan dan titik mana pun di sepanjang lintasan sistem AI, daripada hanya mempertimbangkan posisi akhir. Berdasarkan hal ini, Oracle Success Rate (OSR) menentukan apakah titik mana pun pada lintasan berada dalam ambang batas keberhasilan yang telah ditentukan sebelumnya, menawarkan kriteria keberhasilan yang lebih santai namun informatif.

Ukuran ketepatan jalur. Menilai seberapa akurat sistem AI mengikuti jalur yang diberikan atau jalur ahli, yang sangat penting dalam keadaan di mana rute itu sendiri penting—karena keselamatan, mengikuti instruksi, atau penafsiran. Ukuran penilaian ini biasanya melibatkan perbandingan tingkat urutan antara lintasan sistem AI dan jalur referensi. Jangkauan yang diberi bobot berdasarkan Skor Panjang (CLS) menggabungkan Jangkauan Jalur (PC), yang mengukur seberapa banyak jalur referensi yang dikunjungi, dengan Skor Panjang (LS) yang memberi penghargaan pada lintasan yang panjangnya mirip dengan lintasan ahli. Normalized Dynamic Time Warping (nDTW) memberikan penalti ringan pada penyimpangan dari lintasan referensi sambil memungkinkan keluwesan dalam penyesuaian agar selaras berdasarkan waktu. Keberhasilan yang diberi bobot berdasarkan normalized Dynamic Time Warping (SDTW) membatasi nDTW hanya pada episode yang berhasil, secara bersamaan mencerminkan penyelesaian tugas dan kesetiaan pada jalur yang diharapkan.

Menjawab Pertanyaan Sambil Berada di Lingkungan

Dalam tugas Menjawab Pertanyaan Sambil Berada di Lingkungan (EQA), sistem AI harus secara aktif menavigasi dan mengamati lingkungannya dari perspektif orang pertama untuk mengumpulkan informasi yang diperlukan untuk menjawab pertanyaan tertentu. Proses ini tidak hanya membutuhkan strategi eksplorasi yang efektif tetapi juga kemampuan untuk membuat keputusan yang tepat tentang kapan informasi yang cukup telah dikumpulkan. Sistem AI cerdas harus mampu merencanakan tindakan yang tepat untuk membuat sebesar mungkin perolehan informasi sekaligus mengenali momen optimal untuk berhenti menjelajah dan memberikan respons yang akurat.

Kumpulan data. Melakukan eksperimen menjawab pertanyaan berwujud (*embodied question answering*/EQA) di dunia fisik seringkali dibatasi oleh batasan lingkungan, persyaratan perangkat keras, dan biaya saat digunakan yang tinggi. Untuk mengatasi tantangan ini, program simulasi telah menjadi alat penting untuk mengembangkan dan menilai sistem AI berwujud. Program simulasi menyediakan lingkungan yang terkontrol dan dapat direproduksi, memudahkan pembuatan kumpulan data skala besar, dan memungkinkan pelatihan model yang aman dan hemat biaya. Lingkungan simulasi juga dengan jelas meningkatkan kelayakan penggunaan model terlatih pada sistem AI dunia nyata dengan memungkinkan eksperimen dan penyempurnaan yang cepat. Beberapa kumpulan data EQA patokan uji telah dikembangkan menggunakan platform simulasi, masing-masing dengan karakteristik dan tujuan yang berbeda.

Tabel 1.8 Perbandingan kumpulan data EQA (A-EQA: EQA Aktif; I-EQA: EQA Interaktif; EM-EQA: EQA Ingatan kejadian; OE: Terbuka; MC: Pilihan Ganda; MINOS*: Program simulasi Berdasarkan MINOS)

Kumpulan data	Tahun	Jenis	Sumber data	Program simulasi	Pembuatan kueri	Jawaban	Ukuran
EQA v1	2018	A-EQA	SUNCG	House3D	Berdasarkan Aturan	OE	5,000+
MT-EQA	2019	A-EQA	SUNCG	House3D	Berdasarkan Aturan	OE	19,000+
MP3D-EQA	2019	A-EQA	MP3D	MINOS*	Berdasarkan Aturan	OE	1,136
IQUAD V1	2018	I-EQA	–	AI2THOR	Berdasarkan Aturan	MC	75,000+
VideoNavQA	2019	EM-EQA	SUNCG	House3D	Berdasarkan Aturan	OE	101,000
SQA3D	2022	Hanya QA	ScanNet	–	Manual	MC	33,400
K-EQA	2023	A-EQA	–	AI2THOR	Berdasarkan Aturan	OE	60,000
OpenEQA	2024	A-EQA, EM-EQA	ScanNet, HM3D	Habitat	Manual	OE	1,600+
HM-EQA	2024	A-EQA	HM3D	Habitat	VLM	MC	500
S-EQA	2024	A-EQA	–	VirtualHome	LLM	Biner	–

Tabel 1.8 memberikan gambaran umum tentang kumpulan data representatif. Kumpulan data EQA paling awal, EQA v1, dibuat di program simulasi House3D menggunakan adegan SUNCG buatan. Kumpulan data ini memiliki lebih dari 5.000 pertanyaan di lebih dari 750 lingkungan, dengan empat jenis pertanyaan yang dirumuskan melalui program fungsional. MT-EQA memperluas EQA v1 ke keadaan multi-objek dengan lebih dari 19.000 pertanyaan yang mencakup perbandingan warna, jarak, dan ukuran. MP3D-EQA menggabungkan adegan Matterport3D dalam ruangan yang realistis ke dalam program simulasi MINOS, memungkinkan navigasi yang lebih realistis. IQUAD V1, yang dikembangkan di AI2-THOR, menekankan interaksi dan proses berpikir kegunaan benda, menawarkan lebih dari 75.000 pertanyaan dalam adegan yang diatur sesuai kebutuhan.

VideoNavQA memisahkan navigasi dan proses berpikir dengan menyajikan sistem AI dengan lintasan video. Ia menggunakan House3D untuk menghasilkan lebih dari 100.000 pasangan pertanyaan-video, yang mencakup delapan kategori. SQA3D, berdasarkan ScanNet, memperkenalkan proses berpikir yang lebih rumit pada lebih dari 6.800 adegan dengan 33.400 pertanyaan. K-EQA meningkatkan EQA dengan pengetahuan terstruktur. Dibangun di AI2-THOR, ia menghubungkan entitas ke grafik pengetahuan untuk mendukung 60.000 pertanyaan yang melibatkan proses berpikir logis dan komparatif. OpenEQA berfokus pada kosakata terbuka, keadaan dunia nyata dari ScanNet dan HM3D, mendukung ingatan kejadian dan eksplorasi aktif. Kumpulan data terbaru seperti HM-EQA dan S-EQA memanfaatkan Model AI Bahasa Besar seperti GPT-4 untuk pembangkitan data yang mudah dikembangkan, meningkatkan keragaman dan kualitas. HM-EQA menggunakan adegan HM3D untuk

menghasilkan 500 pertanyaan pilihan ganda, sementara S-EQA menerapkan penyaringan kesamaan kosinus di VirtualHome untuk memastikan relevansi dan variasi. Kumpulan data yang beragam ini menyediakan sumber daya berharga untuk memajukan AI berwujud, mendukung riset tentang pengamatan, proses berpikir, memori, dan eksplorasi di lingkungan 3D simulasi.

Ukuran penilaian. Kinerja dalam menjawab pertanyaan berwujud biasanya dinilai berdasarkan dua aspek utama: navigasi dan menjawab pertanyaan. Untuk navigasi, banyak riset mengikuti cara riset yang diusulkan oleh Das dan kawan-kawan., menggunakan indikator seperti jarak akhir ke objek target (d_T), perubahan jarak dari posisi awal ke posisi akhir (d_Δ), dan jarak minimum ke target yang dicapai selama episode ($d_{\min}$). Ukuran penilaian ini umumnya diukur setelah sejumlah tindakan tetap, seperti 10, 30, atau 50 langkah. Pendekatan lain menilai navigasi menggunakan indikator seperti panjang lintasan dan skor intersection-over-union (IoU) untuk objek target. Dalam hal menjawab pertanyaan, kinerja biasanya dinilai menggunakan ukuran penilaian seperti peringkat rata-rata (MR) dari jawaban yang benar dalam daftar berperingkat dan ketepatan keseluruhan. Untuk menilai jawaban kosakata terbuka, Majumdar dan kawan-kawan. memperkenalkan ukuran penilaian LLM-Match, yang memanfaatkan Model AI Bahasa Besar untuk menilai kebenaran makna dari jawaban. Mereka juga mengusulkan penggabungan panjang lintasan yang dinormalisasi sebagai faktor pembobot, memungkinkan penilaian gabungan dari kebenaran jawaban dan efisiensi sistem AI. Metode-metode ini bersama-sama memberikan penilaian menyeluruh terhadap kemampuan proses berpikir berwujud oleh model.

Penggenggaman dalam Lingkungan Nyata

Interaksi berwujud meluas melampaui menjawab pertanyaan untuk mencakup pelaksanaan tindakan fisik berdasarkan instruksi manusia, seperti menggenggam dan menempatkan objek. Tugas-tugas ini melibatkan interaksi antara robot, manusia, dan objek, yang membutuhkan kemampuan tergabung dalam pemahaman makna, pengamatan adegan, pengambilan keputusan, dan perencanaan kendali. Pendekatan terbaru menggabungkan penggenggaman robot tradisional dengan LLM dan model dasar bahasa-penglihatan, memungkinkan robot untuk menafsirkan dan melakukan tugas di bawah masukan dari sensor dengan banyak jenis data, termasuk Pengamatan Aktif berdasarkan penglihatan dan proses berpikir yang dipandu bahasa. Riset saat ini dalam penggenggaman robot terutama berfokus pada dua jenis efektor akhir: penjepit paralel dua jari dan tangan cekatan lima jari. Untuk penjepit paralel, postur genggaman biasanya digambarkan dalam format 4-DOF atau 6-DOF. Susunan 4-DOF mendefinisikan genggaman menggunakan posisi 3D dan sudut yaw, yang sering disebut sebagai genggaman dari atas ke bawah. Sebaliknya, susunan 6-DOF menentukan posisi dan orientasi 3D penuh untuk genggaman yang lebih luwes. Untuk tangan yang cekatan, seperti ShadowHand dengan 26 jumlah arah gerak, peningkatan dimensi menimbulkan tantangan besar baik untuk sintesis genggaman maupun perencanaan lintasan.

Kumpulan data. Dalam beberapa tahun terakhir, banyak kumpulan data genggaman telah diperkenalkan untuk memajukan riset dalam penggerakan benda oleh robot. Kumpulan data ini biasanya menyediakan susunan genggaman yang diberi label penjelas berdasarkan

berbagai jenis data masukan, seperti gambar RGB, peta kedalaman, kumpulan titik 3D, atau pembentukan ulang adegan 3D penuh. Dengan penggabungan model pembelajaran dengan banyak jenis data dan penggunaan model bahasa dasar dalam pengambilan objek oleh robot, terdapat kebutuhan yang semakin meningkat akan kumpulan data yang menggabungkan informasi linguistik. Sebagai tanggapan, beberapa kumpulan data yang ada telah diperluas atau direstrukturisasi menjadi kumpulan data pengambilan objek makna, yang mendukung pengembangan model yang dapat memahami dan melaksanakan pengambilan objek yang dipandu oleh bahasa. Kumpulan data pengambilan objek tradisional mencakup keadaan objek tunggal dan keadaan adegan yang rumit. Misalnya, kumpulan data seperti menyediakan label penjelas pengambilan objek stabil 4-DOF atau 6-DOF di lingkungan desktop nyata, bersama dengan data RGB, kedalaman, dan kumpulan titik 3D. Yang lain, seperti, dikumpulkan di lingkungan simulasi dan menawarkan model adegan skala besar dan beragam serta label penjelas pengambilan objek.

Meskipun kumpulan data ini sangat cocok untuk melatih langkah kerja pengambilan objek, mereka seringkali kurang memiliki kaitan dengan makna, sehingga membatasi kegunaannya untuk tugas yang dipandu oleh bahasa. Untuk mengatasi keterbatasan ini, beberapa upaya telah dilakukan untuk memperkaya kumpulan data genggaman dengan label penjelas tekstual atau label berdasarkan instruksi. Kumpulan data yang diperkaya ini memungkinkan sistem AI untuk menyelaraskan interaksi gambar dan fisik dengan pemahaman bahasa, sehingga mendorong perilaku genggaman yang lebih cerdas dan sadar konteks. Tabel 1.9 merangkum kumpulan data genggaman biasa dan yang diperkaya bahasa, menekankan fitur-fitur utama dan kontribusinya terhadap bidang ini.

Tabel 1.9 Kumpulan data genggaman berwujud

Kumpulan data	Tahun	Jenis	Jenis data	Label genggaman	Jari gripper	Objek	Genggaman	Adegan	Bahasa
Cornell	2011	Nyata	RGB-D	Rect.	2	240	8K	Tunggal	×
Jacquard	2018	Sim	RGB-D	Rect.	2	11K	1.1M	Tunggal	×
GraspNet	2019	Sim	3D	6D	2	206	7.07M	Tunggal	×
ACRONYM	2021	Sim	3D	6D	2	8872	17.7M	Multi	×
MultiGripperGrasp	2024	Sim	3D	-	2-5	345	30.4M	Tunggal	×
OCID-Grasp	2021	Nyata	RGB-D	Rect.	2	89	75K	Multi	×
OCID-VLG	2023	Nyata	RGB-D, 3D	Rect.	2	89	75K	Multi	✓
ReasoingGrasp	2024	Nyata	RGB-D	6D	2	64	99.3M	Multi	✓
CapGrasp	2024	Sim	3D	-	5	1.8K	50K	Tunggal	✓

Ukuran penilaian. Dalam penggenggaman berwujud, ukuran penilaian mencakup pelaksanaan fisik, keselarasan makna, dan kinerja tingkat tugas. Ukuran penilaian pelaksanaan fisik terutama menilai apakah penggenggaman berhasil secara mekanis. Yang paling umum adalah tingkat keberhasilan penggenggaman, yang didefinisikan sebagai rasio penggenggaman yang berhasil terhadap total percobaan. Untuk metode yang memperkirakan posisi penggenggaman, ketepatan dan recall digunakan untuk mengukur ketepatan deteksi,

sementara kesalahan posisi penggenggaman mengukur penyimpangan antara posisi yang diperkirakan dan posisi sebenarnya menggunakan perbedaan posisi dan rotasi. Ukuran penilaian keselarasan makna menilai apakah perilaku penggenggaman selaras dengan maksud linguistik. Ketepatan penggenggaman makna mengukur proporsi penggenggaman yang secara benar memenuhi instruksi bahasa (misalnya, menggenggam "cangkir merah").

Selain itu, skor dasar bahasa mengukur keselarasan antara bentuk angka ringkas bahasa dan tindakan menggunakan fungsi kesamaan seperti kesamaan kosinus, yang mencerminkan sejauh mana penggenggaman sesuai dengan makna yang dimaksud. Ukuran penilaian tingkat tugas mempertimbangkan apakah penggenggaman memungkinkan penyelesaian tugas hilir yang sukses. Tingkat penyelesaian tugas menghitung rasio tugas yang diselesaikan terhadap tugas yang dicoba, relevan untuk interaksi beberapa langkah seperti pengambilan dan penempatan atau penyortiran. Efisiensi genggaman mengukur biaya berdasarkan waktu atau perilaku per genggaman yang berhasil, dinyatakan sebagai kebalikan dari waktu penyelesaian atau jumlah langkah tindakan. Bersama-sama, ukuran penilaian ini memberikan pandangan holistik tentang kinerja genggaman berwujud, yang tidak hanya mencakup kelayakan mekanis tetapi juga pemahaman makna sistem AI dan efektivitas tugas. Menggabungkan ketiga jenis ukuran penilaian ini sangat penting untuk menilai sistem genggaman modern yang menggabungkan kemampuan penglihatan, bahasa, dan pengambilan keputusan dalam keadaan dunia nyata yang rumit.

BAB 2
AI BERWUJUD OTONOM

2.1 PENDAHULUAN

Agen berwujud yang beroperasi dalam lingkungan yang kompleks dan terbuka memerlukan otonomi yang melampaui tugas yang telah ditentukan sebelumnya dan arahan manusia. Sistem tradisional, meskipun efektif dalam kondisi terkontrol, sering kali gagal ketika menghadapi ketidakpastian, ambiguitas, lingkungan dinamis, atau tujuan yang berkembang. Untuk mencapai pembelajaran otonom sebagaimana bayi belajar, AI Berwujud Otonom mengintegrasikan tiga kapasitas inti: Pembelajaran Lingkungan Otonom, yang memungkinkan agen mengeksplorasi, mempersepsikan, dan menafsirkan lingkungan dinamis dan terbuka secara otonom; Pembelajaran Kognitif Otonom, yang berfokus pada pengembangan otonom keterampilan penalaran dan pengambilan keputusan; serta Interaksi Dinamis Otonom, yang memfasilitasi agen yang berlandaskan pembelajaran lingkungan dan model kognitif untuk terlibat dalam interaksi dinamis otonom dengan entitas (siapa/apa) pada waktu yang tepat (kapan). Kapasitas-kapasitas ini bukan modul independen, melainkan proses yang saling terhubung dan membentuk lingkaran berkelanjutan antara persepsi, kognisi, dan interaksi. Secara bersama-sama, ketiganya memungkinkan agen mengurangi ketidakpastian, memperluas kemungkinan tindakan, dan menghasilkan perilaku bermakna tanpa instruksi eksternal. Kerangka kerja ini menyediakan landasan untuk memajukan AI terwujud dalam domain seperti robotika, mengemudi otonom, dan penemuan ilmiah, ketika kemampuan adaptasi jangka panjang dan pengambilan keputusan independen menjadi esensial.

Seiring semakin banyaknya agen berwujud dikerahkan dalam lingkungan yang kompleks dan terbuka, permintaan akan otonomi menjadi sangat penting. Sistem tradisional sering kali bergantung pada tugas yang didefinisikan secara eksternal, kebijakan yang dibuat secara manual, dan sinyal pembelajaran yang dikelola manusia. Meskipun efektif dalam pengaturan yang terkontrol, sistem tersebut akan goyah ketika dihadapkan dengan situasi yang tidak dikenal, umpan balik yang ambigu, atau tujuan yang berkembang. Dalam domain praktis seperti robotika domestik, mengemudi otonom, dan penemuan ilmiah, agen harus merasakan, berinteraksi dengan, dan bertindak terhadap lingkungan mereka tanpa pengawasan manusia yang konstan. Kebutuhan ini menggarisbawahi pentingnya sistem kecerdasan berwujud yang mampu beradaptasi, memutuskan, dan belajar secara independen. AI berwujud otonom bertumpu pada tiga kapasitas fundamental: Pembelajaran Lingkungan Otonom, Pembelajaran Kognitif Otonom, dan Interaksi Dinamis Otonom. Bersama-sama, kapasitas ini memungkinkan agen untuk belajar dan bertindak secara bermakna di dunia nyata tanpa instruksi eksternal.

- **Pembelajaran Lingkungan Otonom** adalah kapasitas untuk menentukan apa, kapan, dan bagaimana mempersepsikan berdasarkan tujuan internal dan ketidakpastian kontekstual. Tidak seperti penginderaan pasif, hal ini memungkinkan agen untuk mengarahkan perhatian, memodulasi sensor, dan secara selektif memperoleh informasi yang paling relevan dengan keadaan atau tugasnya. Melalui ini, agen

mengurangi ketidakpastian, menyelesaikan ambiguitas, dan mengumpulkan bukti dengan cara yang didorong oleh tujuan.

- **Pembelajaran Kognitif Otonom** berkaitan dengan bagaimana agen berinteraksi dengan lingkungannya untuk memengaruhi keadaan fisik dan informasional. Ini termasuk eksplorasi strategis, penggunaan alat, komunikasi dengan manusia atau agen lain, dan manipulasi lingkungan. Alih-alih bergantung pada aturan yang telah ditentukan sebelumnya, agen belajar kapan dan bagaimana berinteraksi dengan cara yang meningkatkan pemahaman dan memperluas repertoar tindakan yang mungkin dilakukannya.
- **Interaksi Dinamis Otonom** melibatkan pemilihan dan pelaksanaan tindakan berdasarkan tujuan internal, model yang dipelajari, dan dinamika lingkungan. Hal ini memungkinkan perilaku yang tidak hanya reaktif tetapi juga antisipatif dan disengaja. Dengan mengintegrasikan persepsi, memori, dan riwayat interaksi, agen menghasilkan rencana dan tindakan yang koheren dan peka terhadap konteks.

Ketiga kapasitas ini saling terkait erat, bukan terisolasi. Persepsi aktif membentuk informasi apa yang diperoleh, memandu interaksi selanjutnya; interaksi mengubah lingkungan dan keadaan internal agen, memengaruhi keputusan di masa mendatang; dan tindakan mengekspresikan pembelajaran sebelumnya dan menghasilkan umpan balik baru. Bersama-sama, mereka membentuk lingkaran berkelanjutan dari persepsi, interaksi, dan adaptasi yang mendasari otonomi berwujud. Bagian selanjutnya dari bab ini meninjau dasar-dasar otonomi dalam pembelajaran manusia dan mesin, menyoroti mekanisme penetapan tujuan, pengaturan diri, dan evaluasi adaptif. Berdasarkan dasar ini, bab ini memperkenalkan komponen inti dari kecerdasan otonom yang terwujud secara rinci, memeriksa mekanisme internal, saling ketergantungan, dan peran mereka dalam mendukung adaptabilitas jangka panjang.

2.2 TEKNOLOGI UTAMA AI BERWUJUD OTONOM

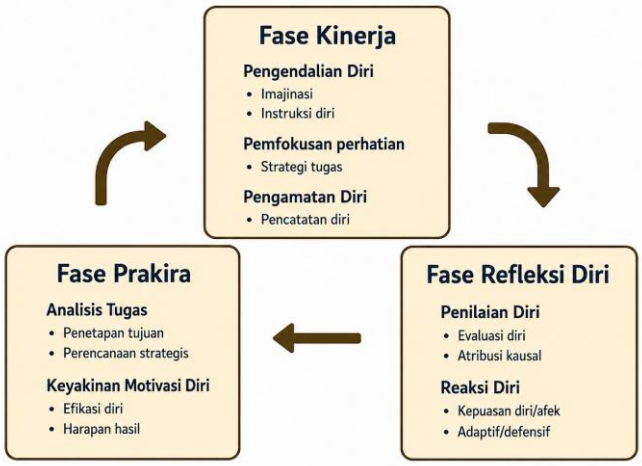
Pembelajaran otonom manusia melibatkan penetapan tujuan, pemilihan strategi, pengaturan perilaku, dan refleksi terhadap hasil. Ini adalah proses dinamis dan peka konteks di mana pembelajar secara proaktif merencanakan, melaksanakan, dan mengevaluasi pembelajaran mereka, menyesuaikan strategi berdasarkan umpan balik dan pengalaman. Proses siklus ini meningkatkan pemecahan masalah, pembelajaran seumur hidup, dan adaptabilitas di lingkungan yang kompleks. Wawasan ini menginformasikan pembelajaran mesin otonom, di mana agen buatan mereplikasi pembelajaran mandiri. Kerangka kerja modern mengimplementasikan modul yang saling bergantung untuk pemilihan tugas, pilihan model, prioritas data, optimasi, dan evaluasi kinerja. Pemilihan tugas menetapkan tujuan, manajemen model dan data memandu pelaksanaan pembelajaran yang efektif, dan evaluasi mendukung penyempurnaan adaptif. Dengan meniru otonomi manusia, sistem otonom mencapai pembelajaran iteratif, berbasis tujuan, dan koreksi diri, memungkinkan perolehan pengetahuan yang efisien, adaptasi yang kuat, dan pemecahan masalah yang efektif dalam domain yang dinamis dan terbuka.

Pembelajaran Otonom pada Manusia

Studi tentang pembelajaran otonom pada manusia telah lama menjadi landasan riset dalam psikologi kognitif, teori pendidikan, dan riset perkembangan. Alih-alih hanya bergantung pada instruksi eksternal, pembelajar otonom mampu menetapkan tujuan mereka sendiri, memilih strategi yang tepat, mengatur perilaku mereka, dan merefleksikan hasil. Kapasitas ini mendasar tidak hanya untuk pembelajaran yang efektif dalam lingkungan terstruktur, tetapi juga untuk pemecahan masalah adaptif, pembelajaran sepanjang hayat, dan eksplorasi kreatif dalam lingkungan terbuka. Di bidang pendidikan orang dewasa, Knowles adalah salah satu sarjana pertama yang mengkonseptualisasikan pembelajaran mandiri. Ia menekankan bahwa pembelajar dewasa mampu secara mandiri mengidentifikasi kebutuhan belajar mereka, merencanakan strategi untuk mengatasinya, dan mengevaluasi kemajuan secara kritis. Perspektif ini, yang menyoroti peran aktif pembelajar, sejak itu telah memberikan pengaruh yang luas di seluruh konteks pendidikan formal dan pembelajaran sepanjang hayat informal.

Teori perkembangan kognitif juga memberikan landasan penting untuk memahami munculnya otonomi dalam pembelajaran. Piaget berpendapat bahwa individu membangun pengetahuan melalui interaksi aktif dengan lingkungan mereka, terus-menerus membentuk kembali struktur kognitif melalui proses asimilasi dan akomodasi. Sebaliknya, Vygotsky menekankan dimensi sosiokultural pembelajaran, dengan berpendapat bahwa otonomi berkembang melalui partisipasi terbimbing dan secara bertahap beralih menuju pengaturan diri independen dalam zona perkembangan proksimal pembelajar. Bidang metakognisi telah lebih jauh memajukan pemahaman ini dengan mengungkapkan bahwa otonomi dalam pembelajaran tidak hanya bersifat perilaku, tetapi juga pada dasarnya reflektif. Flavell memperkenalkan gagasan metakognisi sebagai kapasitas untuk memantau dan mengendalikan proses berpikir sendiri. Berdasarkan hal ini, Schraw dan Moshman membedakan antara pengetahuan deklaratif tentang strategi dan pengetahuan prosedural untuk menerapkannya. Wawasan ini meletakkan dasar bagi model-model kontemporer pembelajaran strategis yang diatur sendiri.

Di antara model-model ini, salah satu kerangka kerja yang paling komprehensif dan berbasis empiris adalah kerangka kerja pembelajaran yang diatur sendiri yang berpengaruh dari Barry Zimmerman. Seperti yang ditunjukkan pada Gambar 2.1, model Zimmerman memandang pembelajaran sebagai proses dinamis dan siklik yang terdiri dari tiga fase yang saling bergantung: perencanaan, pelaksanaan, dan refleksi diri. Setiap fase diatur oleh mekanisme kognitif dan motivasi yang berbeda namun saling melengkapi yang memungkinkan peserta didik untuk merencanakan, secara aktif melaksanakan, dan mengevaluasi perilaku belajar mereka sendiri. Kerangka kerja ini telah divalidasi di berbagai kelompok usia, bidang studi, dan lingkungan pengajaran, dan menawarkan lensa yang kuat untuk meneliti otonomi manusia dalam berbagai konteks pendidikan. Selain itu, kerangka kerja ini berfungsi sebagai jembatan konseptual untuk menganalisis struktur dan fungsi otonomi dalam pembelajaran mesin dan kecerdasan berwujud.



Gambar 2.1 Model siklik tiga fase pembelajaran otonom

Fase Perencanaan Awal

Fase perencanaan awal mendahului perilaku belajar yang dapat diamati dan melibatkan persiapan kognitif dan motivasi untuk keterlibatan tugas yang sukses. Zimmerman menyoroti dua komponen inti dalam fase ini: analisis tugas dan keyakinan motivasi diri. Analisis tugas mencakup kemampuan pembelajar untuk mengevaluasi secara kritis tuntutan tugas dan merumuskan rencana tindakan yang jelas. Ini termasuk menetapkan tujuan spesifik, merancang strategi yang efektif, dan mengantisipasi tantangan potensial. Analisis tugas yang efektif bersifat proaktif dan strategis, di mana pembelajar memanfaatkan pengetahuan sebelumnya, menyelaraskan sumber daya yang tersedia dengan persyaratan tugas, dan menyesuaikan rencana mereka dengan kendala kontekstual. Ini memberikan kerangka struktural penting untuk pembelajaran otonom.

Sama pentingnya adalah keyakinan motivasi diri, yang memengaruhi kemauan pembelajar untuk memulai dan mempertahankan upaya. Ini termasuk efikasi diri, nilai tugas yang dirasakan, hasil yang diantisipasi, dan orientasi tujuan. Pembelajar dengan keyakinan efikasi yang kuat dan penilaian intrinsik terhadap tugas lebih cenderung terlibat secara mendalam dan bertahan melalui kesulitan. Perlu dicatat bahwa keyakinan tersebut bersifat dinamis dan peka terhadap konteks, berkembang seiring pengalaman dan refleksi. Interaksi antara analisis tugas dan motivasi membentuk dasar otonomi, memungkinkan peserta didik untuk mengejar tugas yang sesuai dengan minat dan kemampuan mereka sambil mempertahankan rasa kendali diri.

Fase Kinerja

Pada fase kinerja, peserta didik terlibat dalam pelaksanaan dan pengaturan strategi pembelajaran secara aktif. Fase ini didorong oleh dua proses yang saling terkait: pengendalian diri dan pengamatan diri. Pengendalian diri melibatkan penerapan strategi kognitif secara sengaja (misalnya, elaborasi, latihan, imajinasi) dan pengaturan perhatian, emosi, dan perilaku untuk mengoptimalkan kondisi pembelajaran. Pengelolaan lingkungan juga integral, seperti meminimalkan gangguan atau mengalokasikan waktu secara efektif. Pengaturan diri emosional memainkan peran penting, karena kecemasan atau kepercayaan diri yang berlebihan yang tidak terkendali dapat mengganggu fungsi kognitif. Pengamatan diri

melengkapi pengendalian diri dengan memungkinkan peserta didik untuk memantau kemajuan mereka sendiri dan menyesuaikan perilaku secara waktu nyata. Proses ini melibatkan pelacakan indikator kinerja, mengenali momen kebingungan atau kesalahan, dan menilai efektivitas strategi yang dipilih. Peserta didik dapat menggunakan daftar periksa, protokol berpikir keras, atau catatan kinerja untuk mempertahankan kesadaran berkelanjutan tentang keadaan kognitif mereka. Observasi diri yang efektif memungkinkan koreksi di tengah proses, seperti memperlambat ketika menghadapi kesulitan konseptual atau meninjau kembali konten sebelumnya untuk memperkuat pemahaman. Proses-proses ini memungkinkan bentuk otonomi situasional, di mana peserta didik secara dinamis menyesuaikan perilaku mereka sebagai respons terhadap umpan balik internal dan tuntutan eksternal.

Fase Refleksi Diri

Fase refleksi diri mengikuti penyelesaian tugas dan memberikan kesempatan bagi peserta didik untuk mengevaluasi kinerja mereka dan memperoleh wawasan untuk pembelajaran di masa mendatang. Fase ini terdiri dari dua proses: penilaian diri dan reaksi diri. Penilaian diri melibatkan evaluasi kritis terhadap kinerja seseorang, termasuk apakah tujuan tercapai dan strategi efektif. Peserta didik mempertimbangkan baik hasil maupun penyebab mendasar dari keberhasilan atau kegagalan, dengan mengacu pada atribusi yang terkait dengan upaya, penggunaan strategi, kesulitan tugas, atau kondisi eksternal. Evaluasi ini bukan hanya retrospektif tetapi memainkan peran sentral dalam membentuk motivasi dan perilaku dalam tugas-tugas di masa mendatang.

Reaksi diri mengacu pada respons emosional dan strategis peserta didik terhadap penilaian yang dibentuk. Evaluasi positif dapat mengarah pada kepuasan, peningkatan efikasi diri, dan penguatan strategi yang efektif. Evaluasi negatif, jika diproses secara konstruktif, dapat memotivasi revisi tujuan, pengabaian kebiasaan yang tidak efektif, atau adopsi strategi baru. Yang penting, refleksi juga memfasilitasi abstraksi: peserta didik mulai menggeneralisasi dari pengalaman spesifik dan mengembangkan pemahaman tingkat meta tentang bagaimana mereka belajar dengan terbaik. Ini berpuncak pada identitas pembelajaran yang lebih tangguh dan fleksibel, yang mampu beradaptasi di berbagai konteks dan dari waktu ke waktu.

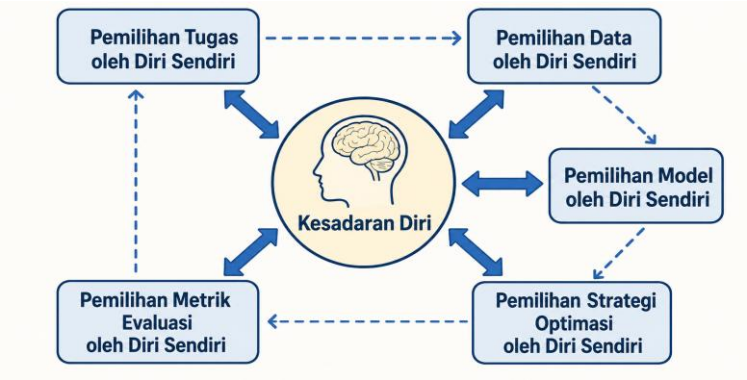
Ketiga fase ini membentuk model pembelajaran otonom yang koheren dan bernuansa. Mereka menyoroti bagaimana para pembelajar secara proaktif memulai, mengelola secara strategis, dan secara reflektif menyempurnakan keterlibatan mereka dengan tugas-tugas pembelajaran. Alih-alih menjadi sifat statis, otonomi diungkapkan di sini sebagai kapasitas dinamis dan peka konteks yang didasarkan pada kesadaran, kemampuan beradaptasi, dan agensi. Kerangka kerja ini tidak hanya memperdalam pemahaman kita tentang pembelajaran manusia tetapi juga memberikan informasi untuk desain sistem pembelajaran otonom pada agen kecerdasan buatan.

Pembelajaran Mesin Otonom

Studi tentang pembelajaran otonom pada manusia memberikan wawasan mendasar untuk pengembangan pembelajaran mesin otonom (AML). Berakar pada psikologi kognitif, metakognisi, dan teori pendidikan, otonomi manusia didefinisikan sebagai kapasitas untuk

menetapkan tujuan, merancang strategi, memantau kemajuan, dan menyesuaikan perilaku berdasarkan evaluasi reflektif. Fungsi kognitif ini tidak hanya mencirikan pembelajaran manusia yang efektif tetapi juga menawarkan kerangka konseptual untuk memungkinkan bentuk otonomi analog dalam sistem cerdas. Dorongan untuk AML muncul dari meningkatnya kompleksitas dan dinamika lingkungan pembelajaran dunia nyata. Alur kerja pembelajaran mesin tradisional membutuhkan pengawasan manusia yang ekstensif untuk definisi tugas, kurasi data, desain model, evaluasi, dan penerapan. Ketergantungan tersebut membatasi skalabilitas, adaptabilitas, dan generalisasi. Pembelajaran mesin otonom bertujuan untuk mengatasi keterbatasan ini dengan memberikan agen buatan kemampuan untuk memandu dan mengoptimalkan perilaku pembelajaran mereka dengan intervensi manusia minimal.

Riset terbaru telah memformalkan pembelajaran mesin otonom sebagai kerangka kerja yang koheren yang bertujuan untuk pembelajaran mandiri di semua tahapan pipeline. Zhu dan kawan-kawan mengusulkan kerangka kerja pembelajaran mesin mandiri, menekankan bahwa AML tidak boleh terbatas pada otomatisasi komponen terisolasi seperti pelatihan model atau penyetelan hiperparameter, tetapi harus mendukung sistem siklus tertutup sepenuhnya yang mampu menemukan tugas, memilih model, mengelola data, mengevaluasi kinerja, dan memperbaiki perilaku pembelajarannya sendiri. Formulasi mereka membingkai AML sebagai masalah optimasi bertingkat, di mana ruang tugas, ruang model, ruang data, dan ruang tujuan semuanya tunduk pada pencarian dan penyempurnaan adaptif. Pendekatan ini memungkinkan pembelajar untuk menentukan tidak hanya solusi optimal, tetapi juga masalah apa yang harus dipecahkan dan bagaimana mendekatinya. Secara lebih konkret, seperti yang ditunjukkan pada Gambar 2.2, proses AML terstruktur menjadi enam modul yang saling bergantung: konstruksi tugas mandiri, pemilihan model mandiri, kurikulum data mandiri, pemilihan metrik evaluasi mandiri, pemilihan strategi optimasi mandiri, dan penerapan mandiri. Modul-modul ini saling terkait secara rekursif: misalnya, konstruksi tugas membentuk pemilihan model; optimasi memengaruhi prioritas data; evaluasi mendorong revisi strategi. Arsitektur berlapis ini mencerminkan sifat rekursif dan adaptif dari pembelajaran manusia.



Gambar 2.2 Kerangka kerja pembelajaran mesin mandiri

Pemilihan Tugas Mandiri

Pemilihan tugas mandiri mengacu pada kemampuan sistem untuk secara otonom menentukan tugas yang paling bermanfaat untuk dikejar dari ruang tugas kandidat. Proses ini

dipandu oleh fungsi utilitas $\mathcal{V}(\mathcal{T})$ yang mengukur nilai yang diharapkan dari setiap tugas, dengan mempertimbangkan hal-hal seperti potensi pembelajaran, kebaruan, kesulitan, atau utilitas jangka panjang. Pembelajar memilih tugas $\mathcal{T}^*$ yang memaksimalkan utilitas ini:

$$\mathcal{T}^* = \underset{\mathcal{T} \in \mathcal{T}}{\operatorname{argmax}} \mathcal{V}(\mathcal{T}) \tag{2.1}$$

Kemampuan ini memungkinkan sistem AML untuk bertindak melampaui pembelajaran reaktif dan terlibat dalam pengambilan keputusan proaktif, mirip dengan bagaimana seorang pembelajar manusia memilih keterampilan mana yang akan dikembangkan atau kesenjangan pengetahuan mana yang akan diatasi. Dalam praktiknya, pemilihan tugas dapat melibatkan pembelajaran embedding tugas, memperkirakan imbalan di masa depan dari penguasaan tugas, atau menggunakan strategi eksplorasi untuk menyeimbangkan tujuan yang diketahui dan tidak diketahui. Yang penting, modul ini meletakkan dasar untuk siklus pembelajaran yang berorientasi pada tujuan, memastikan bahwa lintasan pembelajaran sistem selaras dengan tujuan internal atau eksternal.

Pemilihan Model Mandiri

Setelah tugas diidentifikasi dengan jelas, pembelajar harus dengan cermat memilih arsitektur model atau representasi hipotesis yang tepat untuk berhasil melaksanakan tugas secara efektif. Pemilihan model mandiri melibatkan evaluasi sistematis beberapa model kandidat $m \in \mathcal{M}$ dan memilih model yang menawarkan kinerja generalisasi yang diharapkan terbaik. Secara formal, ini dinyatakan sebagai:

$$m^* = \arg\min_{m \in \mathcal{M}} \mathbb{E}_{(x,y)\sim\mathcal{D}}[\ell(m(x), y)] \tag{2.2}$$

Pemilihan ini dapat diwujudkan melalui pencarian arsitektur neural (NAS), optimasi hiperparameter, atau meta-modeling berdasarkan kinerja model tugas sebelumnya. Dari perspektif fungsional, modul ini memungkinkan pembelajar untuk secara adaptif membangun representasi internal dan bias induktif yang spesifik untuk tugas, sehingga meningkatkan akurasi dan efisiensi sampel. Ini juga mencerminkan perilaku manusia dalam memilih model mental atau perangkat konseptual yang tepat untuk mendekati tantangan pembelajaran.

Pemilihan Data Mandiri

Pemilihan data merupakan bagian penting dari efisiensi pembelajaran. Alih-alih melatih pada dataset statis atau yang diambil secara acak, pembelajar otonom terus menerus memprioritugasan atau menyaring sampel berdasarkan nilai pembelajarannya. Dalam AML, ini dioperasionalkan melalui pemberian bobot w_i pada instance data (x_i, y_i) dan mengoptimalkan kerugian berbobot:

$$\min_{\theta} \sum_{i} w_i \cdot \ell(f_\theta(x_i), y_i) \tag{2.3}$$

Bobot w_i dapat dihitung berdasarkan estimasi ketidakpastian, tujuan pembelajaran kurikulum, atau ukuran keragaman sampel. Mekanisme ini memungkinkan pembelajar untuk membangun kurikulumnya sendiri dengan berfokus pada contoh-contoh yang sulit, mengidentifikasi pola yang kurang terwakili, atau melewatkan sampel yang berlebihan, yang sejajar dengan cara manusia memilih untuk mempelajari materi yang lebih menantang atau meninjau kesalahan sebelumnya. Pemilihan data mandiri yang efektif mengarah pada konvergensi yang lebih cepat dan generalisasi yang lebih kuat, terutama dalam lingkungan yang tidak stasioner atau terbuka.

Pemilihan Strategi Optimasi Mandiri

Selain pemilihan model dan data, pembelajar juga harus menentukan cara pelatihan, yaitu strategi optimasi apa yang akan diterapkan. Ini termasuk keputusan tentang aturan pembaruan gradien, jadwal laju pembelajaran, teknik regularisasi, dan konfigurasi pengoptimal. Secara formal, pembelajar memilih strategi optimal $s \in \mathcal{S}$ dengan meminimalkan meta-objektif atas beberapa tugas:

$$s^* = \underset{s \in \mathcal{S}}{\operatorname{argmin}} \mathbb{E}_{\mathcal{T} \sim p(\mathcal{T})} \left[\mathcal{L}^{\text{val}}(\theta_s^*) \right] \tag{2.4}$$

Modul ini menangkap kapasitas untuk belajar belajar, di mana sistem mengidentifikasi dinamika pelatihan yang paling efektif di berbagai konteks. Teknik seperti pembelajaran penguatan untuk desain pengoptimal, pelatihan berbasis populasi, dan metode meta-gradien termasuk dalam kategori ini. Secara konseptual, ini menyerupai bagaimana pembelajar manusia bereksperimen dengan berbagai kebiasaan belajar atau rutinitas pemecahan masalah dan secara bertahap menginternalisasi kebiasaan yang mengarah pada peningkatan kinerja.

Pemilihan Metrik Evaluasi Diri

Modul terakhir dalam siklus AML berkaitan dengan bagaimana pembelajar mengevaluasi kemajuannya. Alih-alih bergantung pada metrik tetap yang dipilih oleh perancang manusia, sistem otonom harus menentukan kriteria evaluasi mana yang paling baik menangkap keberhasilannya untuk tugas yang diberikan. Dengan ruang metrik potensial $\mathcal{E}$, pembelajar memilih metrik yang memaksimalkan relevansi tugas:

$$e^* = \underset{e \in \mathcal{E}}{\operatorname{argmax}} \mathcal{R}(e, \mathcal{T}) \tag{2.5}$$

Pemilihan ini mungkin bergantung pada sensitivitas statistik, keselarasan dengan tujuan hilir, atau korelasi metrik tugas historis. Misalnya, dalam tugas bahasa, pembelajar dapat memilih BLEU, ROUGE, atau BERTScore berdasarkan seberapa baik metrik tersebut mencerminkan fidelitas semantik untuk kasus penggunaan tertentu. Modul ini memainkan peran kunci dalam memungkinkan kesadaran metakognitif: modul ini memungkinkan sistem untuk menilai kinerja dalam konteks, menyesuaikan strategi evaluasi dari waktu ke waktu, dan menutup lingkaran umpan balik yang diperlukan untuk peningkatan berkelanjutan.

Kerangka kerja lima lapis pembelajaran mesin otonom sangat mirip dengan proses tiga fase pembelajaran mandiri manusia. Pada fase perencanaan, pembelajar manusia terlibat

dalam penetapan tujuan, analisis tugas, dan pembentukan keyakinan motivasi yang menentukan kemauan dan kesiapan mereka untuk belajar. Ini sesuai langsung dengan modul pemilihan tugas mandiri dalam pembelajaran mesin otonom, di mana sistem mengidentifikasi tugas yang paling bermanfaat untuk dikejar dari ruang tugas yang diberikan. Sama seperti pembelajar manusia yang bergantung pada evaluasi internal tentang relevansi, tantangan, atau minat, pembelajar mesin memperkirakan fungsi utilitas atas tugas-tugas kandidat dan memilih tujuan yang memaksimalkan nilai pembelajaran jangka panjang. Fase ini meletakkan dasar kognitif untuk semua perilaku selanjutnya baik pada manusia maupun mesin.

Fase kinerja dalam pembelajaran manusia melibatkan implementasi strategi, pemantauan waktu nyata, dan kontrol adaptif atas proses kognitif dan perilaku. Dalam pembelajaran mesin otonom, tahap ini tercermin dalam tiga modul yang saling terkait: pemilihan model mandiri, pemilihan data mandiri, dan pemilihan strategi optimasi mandiri. Modul-modul ini mengatur bagaimana sistem memilih arsitektur model yang paling sesuai, memilih dan memberi bobot pada data pelatihan yang informatif, dan menyesuaikan strategi pembelajarannya untuk memastikan konvergensi yang efisien. Bersama-sama, mereka memungkinkan sistem untuk secara dinamis mengatur proses pembelajarannya, seperti halnya pembelajar manusia yang menyempurnakan pendekatan mereka sebagai respons terhadap kemajuan dan kesulitan.

Terakhir, fase refleksi diri, di mana pembelajar manusia mengevaluasi kinerja mereka dan merevisi rencana masa depan mereka, sesuai dengan modul pemilihan metrik evaluasi mandiri. Di sini, mesin memilih kriteria evaluasi yang tepat yang paling mencerminkan keberhasilan dalam tugas saat ini, memungkinkannya untuk menyempurnakan mekanisme penilaian internalnya dan memandu peningkatan di masa mendatang. Keselarasan struktural antara otonomi manusia dan mesin ini menggambarkan logika bersama dari pembelajaran iteratif, berorientasi tujuan, dan koreksi diri.

2.3 KOMPONEN UTAMA AI BERWUJUD OTONOM

Sistem otonom yang terwujud mengintegrasikan persepsi, kognisi, dan tindakan untuk beroperasi secara independen di lingkungan yang kompleks dan dinamis. Inti dari kemampuan ini adalah pengembangan representasi adegan yang kaya, yang memungkinkan robot untuk memahami struktur spasial, properti objek, dan hubungan antar objek. Teknik seperti Neural Radiance Fields (NeRF) memfasilitasi rekonstruksi 3D dengan fidelitas tinggi dari data visual yang jarang, meningkatkan persepsi, lokalisasi, dan pengambilan keputusan. Melengkapi kemajuan ini, Model Bahasa Besar (*Large Language Models*/LLM) dan model bahasa-visual menyediakan penalaran tingkat tinggi, perencanaan, dan landasan multimodal, memungkinkan robot untuk menafsirkan instruksi, menghasilkan rencana yang dapat ditindaklanjuti, dan beradaptasi dengan skenario baru. Model multimodal yang terwujud memperluas kemampuan ini dengan mengintegrasikan penalaran spasial, pengambilan keputusan berurutan, dan kontrol waktu nyata, menjembatani persepsi dan tindakan. Dipadukan dengan perwujudan perangkat keras otonom dan komputasi neuromorfik, sistem ini mencapai interaksi latensi rendah dan hemat energi dengan dunia fisik. Bersama-sama,

perkembangan ini membentuk fondasi untuk robot yang tangguh, fleksibel, dan mampu secara kognitif, memajukan batas-batas pembelajaran otonom, interaksi dinamis, dan perilaku terwujud yang cerdas.

Pembelajaran Lingkungan Otonom

Pembelajaran lingkungan otonom memungkinkan agen untuk secara aktif menjelajahi dan mempersepsikan lingkungan sekitarnya dan secara otonom memperbarui kognisi lingkungan. Pendekatan pembelajaran lingkungan otonom dapat dibagi menjadi tiga kategori: pendekatan berbasis NeRF, berbasis model dasar, dan berbasis LLM.

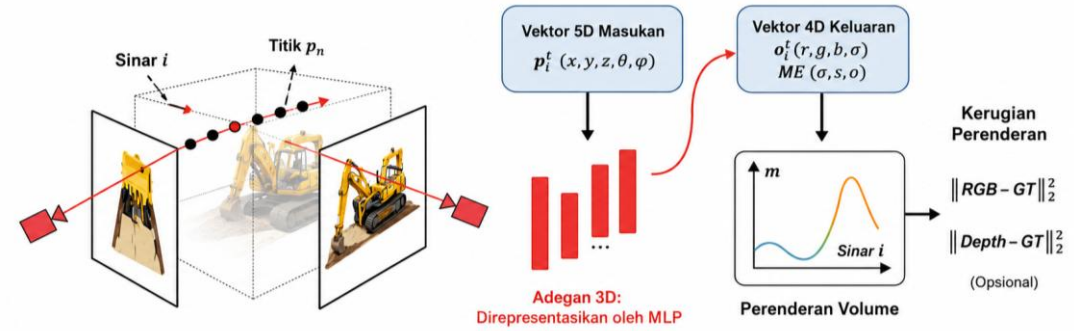
Berbasis NeRF

Tujuan utama representasi adegan adalah untuk mengkodekan lingkungan yang relevan dengan tugas ke dalam format terstruktur yang dapat disimpan dan diproses secara efisien oleh sistem komputasi. Representasi adegan yang baik harus menangkap tidak hanya atribut objek individual, seperti bentuk, ukuran, posisi, dan interaktivitas, tetapi juga hubungan spasial antar objek, termasuk jarak dan oklusi. Representasi tersebut memungkinkan sistem otonom untuk secara akurat memahami struktur spasial lingkungan sekitarnya, melacak perubahan dinamis di lingkungan, dan mendukung penalaran dan pengambilan keputusan yang efektif. Rekonstruksi adegan tiga dimensi mengacu pada proses membangun model 3D digital dari suatu adegan berdasarkan data yang tersedia, seperti gambar, video, atau awan titik. Dibandingkan dengan citra atau video 2D tradisional, rekonstruksi 3D memberikan deskripsi lingkungan yang lebih komprehensif dan sadar geometri, menawarkan dasar yang lebih kaya untuk representasi adegan. Perpaduan kecerdasan buatan dan robotika otonom, khususnya melalui teknik pembelajaran mendalam, telah membawa kemajuan transformatif dalam cara robot memahami dan berinteraksi dengan lingkungan sekitarnya. Di antara berbagai teknologi yang berkontribusi pada kemajuan ini, NeRF telah muncul sebagai metode yang sangat menjanjikan dan berpengaruh. Awalnya diusulkan oleh Mildenhall dan kawan-kawan, NeRF memperkenalkan paradigma baru untuk sintesis tampilan baru dengan menggunakan jaringan saraf untuk memodelkan struktur 3D suatu adegan.

Secara khusus, NeRF menggunakan MLP untuk memetakan input lima dimensi, yang terdiri dari koordinat spasial tiga dimensi dan arah pandang dua dimensi, ke kepadatan volume dan warna RGB. Dilatih pada kumpulan gambar yang jarang, pemetaan ini memungkinkan pembuatan tampilan baru fotorealistik dengan fidelitas yang luar biasa. Seiring waktu, kegunaan NeRF telah meluas ke berbagai domain riset. Di bidang robotika otonom, NeRF menjadi sangat berharga untuk meningkatkan kemampuan persepsi dan memungkinkan interaksi yang lebih canggih antara robot dan lingkungannya. Teknik persepsi robot konvensional sering mengandalkan peta kedalaman, awan titik, atau data sensor serupa untuk merekonstruksi lingkungan sekitar dan mendukung pengambilan keputusan. Meskipun efektif dalam skenario terkontrol, pendekatan ini cenderung menurun secara signifikan ketika data masukan jarang atau bising, atau ketika robot beroperasi di lingkungan yang dinamis dan tidak dapat diprediksi. NeRF menawarkan alternatif dengan merepresentasikan adegan 3D melalui

jaringan saraf, yang memungkinkan rekonstruksi model 3D berkualitas tinggi bahkan dari data yang terbatas.

Robot otonom adalah sistem yang sangat kompleks yang terdiri dari banyak komponen dan subsistem yang saling berinteraksi. Selama beberapa dekade terakhir, berbagai arsitektur robot telah dikembangkan. Untuk kejelasan, sistem ini dapat diabstraksikan menjadi enam modul fundamental: persepsi, lokalisasi dan navigasi, pengambilan keputusan, eksekusi, komunikasi, dan pemantauan keadaan dengan deteksi dan penanganan kesalahan. NeRF telah menunjukkan efektivitas khusus dalam tiga modul pertama, secara konsisten memberikan peningkatan yang signifikan dibandingkan teknik konvensional dalam robotika. Di bidang persepsi, NeRF memungkinkan rekonstruksi objek kaku dan lingkungan skala besar yang sangat akurat, yang sangat penting untuk tugas-tugas robot seperti manipulasi, perakitan, dan inspeksi. Selain itu, NeRF juga mendukung pemodelan objek yang dapat berubah bentuk dan adegan dinamis, sehingga secara signifikan meningkatkan kemampuan adaptasi dan ketahanan secara keseluruhan. Lebih lanjut, NeRF berkontribusi pada pemahaman adegan semantik tingkat lanjut dengan memfasilitasi segmentasi yang lebih detail, memungkinkan robot untuk membedakan antara beberapa objek dan permukaan dengan presisi tinggi.



Gambar 2.3 Gambaran umum kerangka kerja NeRF

Dalam lokalisasi dan navigasi, NeRF memainkan peran penting dalam estimasi pose kamera yang presisi dan lokalisasi dan pemetaan simultan (SLAM). Kemampuannya yang unik untuk mensintesis geometri 3D dengan fidelitas tinggi dari tampilan yang jarang membantu mengatasi tantangan dalam estimasi pose, terutama ketika data pelatihan yang tersedia terbatas. NeRF juga meningkatkan sistem SLAM dengan menyediakan rendering informasi RGB dan kedalaman yang dapat dibedakan, yang mendukung optimasi iteratif pose kamera dan struktur adegan yang detail. Kemampuan canggih ini memungkinkan robot untuk membangun peta yang lebih akurat dan merencanakan jalur optimal secara lebih efektif di lingkungan yang kompleks dan dinamis. Dalam pengambilan keputusan, NeRF mendukung SLAM aktif, memungkinkan robot untuk menjelajahi dan memetakan lingkungan yang tidak dikenal secara waktu nyata. Fungsi ini sangat penting untuk robot bergerak seperti drone dan kendaraan otonom yang beroperasi di lingkungan dinamis yang tidak terstruktur. Selain itu, representasi adegan yang kaya yang disediakan oleh NeRF memfasilitasi proses pengambilan keputusan

berbasis pembelajaran, termasuk pengambilan objek dan optimasi kebijakan, memungkinkan robot untuk lebih memahami dan berinteraksi dengan lingkungan fisik di sekitarnya.

*Representasi adegan***.** NeRF merepresentasikan adegan 3D secara implisit melalui parameter jaringan saraf. Secara khusus, ia menggunakan perceptron multi-layer (MLP) untuk memetakan koordinat spasial ke struktur dasar adegan. Seperti yang ditunjukkan pada Gambar 2.3, kerangka kerja NeRF asli mengambil input 5D $(\boldsymbol{x}, \boldsymbol{d})$, di mana $\boldsymbol{x} = (x, y, z)^T$ menunjukkan posisi spasial 3D dan $d = (\theta, \phi)^T$ menunjukkan arah pandang 2D. Jaringan tersebut menghasilkan nilai kepadatan volumetrik σ_i untuk mengkodekan geometri permukaan secara implisit. Selain kepadatan volume, karya-karya selanjutnya telah mengeksplorasi representasi implisit lainnya seperti fungsi jarak bertanda (SDF) (terpotong) s_i dan probabilitas okupansi o_i. Bersama dengan warna RGB $\boldsymbol{c} = (r, g, b)^T$, keluaran MLP akhir adalah vektor 4D $(\boldsymbol{c}, m)$, di mana m mewakili salah satu dari $[\sigma, s, o]$. Misalkan Θ menyatakan parameter jaringan, fungsi pemetaan NeRF dirumuskan sebagai:

$$\boldsymbol{F}_{\Theta}(\boldsymbol{x}, \boldsymbol{d}) = (\mathbf{c}, m) \tag{2.6}$$

Pengkodean posisi. Model NeRF mengamati bahwa memasukkan input 5D mentah $(\boldsymbol{x}, \boldsymbol{d})$ secara langsung ke MLP menyebabkan rekonstruksi detail frekuensi tinggi yang buruk, sehingga menghasilkan rendering yang kabur. Untuk mengatasi masalah ini, pengkodean posisi diterapkan untuk mengangkat input ke ruang berdimensi lebih tinggi. NeRF asli mengadopsi skema pengkodean sinusoidal yang didefinisikan sebagai:

$$\gamma(\boldsymbol{p}) = \left(\sin(2^0\pi\boldsymbol{p}), \cos(2^0\pi\boldsymbol{p}), \dots, \sin(2^{L-1}\pi\boldsymbol{p}), \cos(2^{L-1}\pi\boldsymbol{p})\right) \tag{2.7}$$

di mana $\boldsymbol{p}$ sesuai dengan $\boldsymbol{x}$ atau $\boldsymbol{d}$. Sebuah studi lanjutan memperkenalkan pemetaan fitur Fourier:

$$\gamma(\boldsymbol{p}) = \left(\sin(2\pi\, \boldsymbol{B}\, \boldsymbol{p}), \cos(2\pi\, \boldsymbol{B}\, \boldsymbol{p})\right) \tag{2.8}$$

Di sini, setiap entri matriks $\boldsymbol{B} \in \mathbb{R}^{L \times d}$ diambil dari distribusi Gaussian dengan rata-rata nol, dengan kovariansi berfungsi sebagai hiperparameter yang dapat disesuaikan dan dipilih dengan cermat berdasarkan tugas spesifik. **Rendering volume.** Untuk memulihkan geometri dan penampilan di sepanjang sinar kamera, teknik rendering volume klasik digunakan. Diberikan sinar $\boldsymbol{r}(t) = \boldsymbol{o} + t\boldsymbol{d}$, di mana o adalah titik asal kamera dan $t \in [t_n, t_f]$ mewakili batas dekat dan jauh, NeRF asli mengambil sampel N titik t_i secara seragam di sepanjang segmen sinar. Estimasi warna sinar $\hat{C}(\boldsymbol{r})$ dihitung dengan:

$$\hat{C}(\mathbf{r}) = \sum_{i=1}^{N} T_i \left(1 - \exp(-\sigma_i \delta_i)\right)\mathbf{c}_i, \quad \text{where} \quad T_i = \exp\left(-\sum_{j=1}^{i-1} \sigma_j\, \delta_j\right) \tag{2.9}$$

Dalam persamaan ini, $\delta_i = t_{i+1} - t_i$ menunjukkan interval antara sampel berurutan. Saat menggunakan fungsi jarak bertanda seperti pada, estimasi warna:

$$\hat{C}(\mathbf{r}) = \frac{1}{\sum_{i=1}^{N} w_i} \sum_{i=1}^{N} w_i \, \mathbf{c}_i, \quad \text{where} \quad w_i = \sigma\left(\frac{s_i}{tr}\right) \cdot \sigma\left(-\frac{s_i}{tr}\right) \tag{2.10}$$

Di sini, $\sigma(\cdot)$ menunjukkan fungsi sigmoid, dan $t\,r$ adalah ambang pemotongan khusus tugas. Terakhir, ketika menggunakan okupansi sebagai representasi implisit, warna sinar didekati menggunakan:

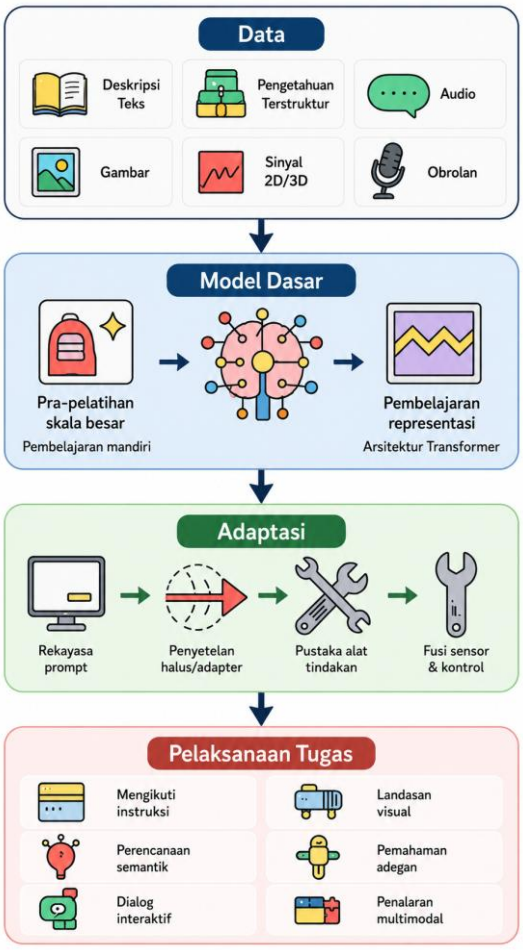
$$\hat{C}(\mathbf{r}) = \sum_{i=1}^{N} w_i \, \mathbf{c}_i, \quad \text{where} \quad w_i = o_i \prod_{j=1}^{i-1} (1 - o_j) \tag{2.11}$$

Formulasi ini memodelkan transmisi sepanjang sinar sebagai hasil kali probabilitas tidak ditempati hingga setiap titik.

Model Berbasis Fondasi

Sistem robot semakin membutuhkan kemampuan kognitif tingkat tinggi, seperti menafsirkan maksud pengguna, menghasilkan rencana yang dapat dieksekusi, dan beradaptasi dengan lingkungan dinamis. Tuntutan ini telah menempatkan model bahasa besar (LLM) sebagai penggerak utama kecerdasan robot tingkat lanjut. Kemampuan mereka untuk mendukung penalaran umum, dapat ditafsirkan, dan sadar konteks telah berkembang seiring dengan evolusi arsitektur dan strategi pelatihan mereka. Penskalaan desain berbasis transformer dan pelatihan pada kumpulan data yang besar dan beragam telah menghasilkan kemampuan baru termasuk mengikuti instruksi, inferensi akal sehat, dan pembangkitan kode. Kemampuan tersebut membuka peluang baru dalam robotika, mulai dari perencanaan bahasa alami hingga landasan multimodal dan kontrol perilaku interaktif.

Kemajuan terbaru semakin memperluas kegunaan LLM untuk robot. Peningkatan penalaran, peningkatan manajemen dialog, dan dukungan untuk input visual memperluas penerapannya dalam pengambilan keputusan berbasis persepsi. Secara paralel, model sumber terbuka dan dapat disesuaikan menawarkan penerapan pada perangkat keras lokal, faktor penting untuk sistem waktu nyata yang sensitif terhadap privasi. Menelusuri perkembangan dari tonggak awal hingga model canggih saat ini memberikan dasar untuk memahami bagaimana LLM diadaptasi dan disematkan ke dalam arsitektur yang terwujud (lihat Gambar 2.4). Meskipun LLM unggul dalam memproses dan menghasilkan teks, sistem robot sering mengandalkan persepsi multimodal untuk bertindak secara efektif di dunia fisik. Model bahasa-visi (VLM) memperluas LLM dengan kemampuan pemahaman visual, memungkinkan penalaran bersama atas bahasa dan citra. Model-model ini mendukung tugas-tugas seperti pemberian keterangan gambar, menjawab pertanyaan visual (VQA), dan pemahaman ekspresi rujukan, yang semuanya penting untuk mendasarkan perintah linguistik dalam konteks dunia nyata.

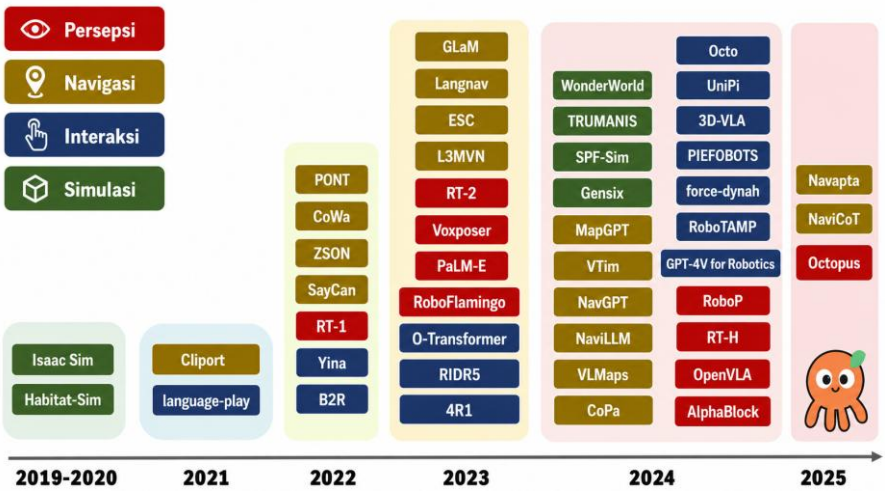


Gambar 2.4 Model dasar untuk mempelajari representasi tujuan umum

Pengembangan VLM telah berkembang dari arsitektur gabungan awal yang menyelaraskan fitur visual dan tekstual hingga pendekatan pembelajaran kontrastif skala besar, seperti CLIP, yang memungkinkan pengenalan kosakata terbuka melalui penyematan gambar-teks bersama. Hal ini memungkinkan robot untuk mengidentifikasi dan menemukan objek tanpa pelatihan kategori eksplisit, memfasilitasi tugas-tugas seperti pemetaan semantik dan pengambilan dari petunjuk bahasa alami. Skalabilitas arsitektur transformer multimodal, seperti yang ditunjukkan oleh SimVLM dan Flamingo, lebih lanjut menunjukkan bahwa penalaran visual few-shot, keterangan adegan, dan dialog dapat dicapai tanpa penyempurnaan khusus tugas, menyediakan antarmuka persepsi serbaguna untuk agen yang berwujud. Pendekatan modular terbaru, yang dicontohkan oleh BLIP-2, menghubungkan encoder gambar beku dengan LLM melalui adaptor ringan, mencapai kinerja visual yang kuat dengan biaya pelatihan yang lebih rendah. Desain semacam itu memungkinkan pemahaman adegan secara waktu nyata, mengikuti instruksi secara responsif, dan pengambilan keputusan yang berlandaskan konteks, sehingga praktis untuk diterapkan dalam sistem robotika. Secara keseluruhan, VLM mewakili langkah penting menuju kecerdasan multimodal yang terukur dan efisien dalam penggunaan sampel, menjembatani kesenjangan antara masukan sensorik mentah dan kognisi berbasis bahasa dalam AI berwujud.

Berbasis LLM

Tidak seperti jaringan saraf tradisional atau model skala besar yang berfokus terutama pada pengenalan objek, Model Bahasa Multimodal yang Terwujud (EMLM) dirancang sesuai dengan konsep kecerdasan berwujud, di mana agen dapat merasakan, berinteraksi dengan, dan menavigasi di dalam dunia fisik. Mencapai kemampuan ini membutuhkan EMLM untuk memiliki kemampuan persepsi dan penalaran tingkat lanjut yang melampaui pengenalan objek hingga pemahaman hubungan spasial dan ketergantungan logis di antara mereka. Dalam konteks ini, persepsi yang terwujud mencakup tidak hanya pengenalan visual tetapi juga penalaran spasial, pemahaman struktur adegan tiga dimensi, dan kemampuan untuk merencanakan dan melaksanakan tugas-tugas kompleks langsung dari masukan visual. Lintasan pengembangan model persepsi yang terwujud diilustrasikan pada Gambar 2.5. Riset terkini tentang model besar untuk persepsi yang terwujud dapat dikategorikan secara luas menjadi dua jenis utama: model yang dibangun di atas arsitektur GPT dan model yang dikembangkan menggunakan arsitektur skala besar alternatif, seperti yang dirangkum dalam Tabel 2.1.



Gambar 2.5 Garis waktu persepsi, navigasi, dan interaksi yang terwujud

Tabel 2.1 Model besar persepsi berwujud

Model	Arsitektur	Platform	Dataset	Tahun
Octopus	MPT-7B	OctoGibson, OctoGTA	–	2025
AlphaBlock	Vicuna, ViT-G/14	Lengan robot Franka Research 3	AlphaBlock	2023
CoPa	GraspNet	Franka Emika Panda	–	2024
ReKep	DINOv2, GPT-4o	Lengan Tunggal Beroda, Lengan Ganda Stasioner	–	2024
RobotGPT	–	PyBullet, MoveIt, ROS Franka	–	2024
Voxposer	–	Franka Emika Panda	SAPIEN	2023
PaLM-E	ViT, OSRT	TAMP, Language-Table, Manipulasi Bergerak	–	2023

RT-1	EfficientNet-B3	Everyday Robots	RT-1	2022
RT-2	PaLI-X, PaLM-E	Manipulator Bergerak 7DoF	–	2023
RT-H	ViT, PaLI-X	–	–	2024
RoboFlamingo	OpenFlamingo	Franka Emika Panda	CALVIN	2023
Openvla	Prismatic-7B	Robot WidowX	BridgeData V2	2024
RoboMamba	Mamba	Franka Emika Panda, SAPIEN	LLaVA-LCS, LLaVA-v1.5, LRV-INSTRUCT, RoboVQA	2024

Model berbasis GPT mampu menginterpretasikan instruksi bahasa alami spesifik tugas dan menghasilkan keluaran pemahaman adegan dalam bentuk teks, sehingga memungkinkan mereka untuk mendukung berbagai tugas persepsi. Contoh representatif termasuk GPT-1, GPT-2, GPT-3, dan GPT-4. Misalnya, Octopus menggunakan GPT-4V untuk secara dinamis mendeskripsikan dan menganalisis gambar yang diamati pada berbagai tahap tugas, mengidentifikasi objek yang dapat berinteraksi dan posisi spasialnya, yang kemudian digunakan sebagai masukan untuk pengambilan keputusan. Demikian pula, CoPa menerapkan GPT-4V untuk mendeteksi dan menyoroti wilayah genggaman spesifik dan potensi pose genggaman, sehingga memungkinkan pemahaman objek yang lebih rinci, sementara GraspNet memilih pose genggaman yang paling tepat untuk dieksekusi.

Meskipun GPT-4V menunjukkan kemampuan yang kuat dalam penalaran multimodal, ia sering kali menghadapi keterbatasan dalam menangani input berbasis koordinat karena persepsi spasial yang terbatas. Untuk mengurangi tantangan ini, AlphaBlock, yang mengambil inspirasi dari penalaran rantai pemikiran, mengharuskan GPT-4V untuk terlebih dahulu menalar tentang tata letak spasial sebelum menghasilkan prediksi koordinat. Hal ini menghasilkan informasi deskriptif dan posisional, yang kemudian dimanfaatkan untuk memberikan umpan balik eksekusi secara waktu nyata. Pendekatan lain mengintegrasikan ekstraksi fitur visual dengan model bahasa, seperti MiniGPT-4, yang menggunakan adaptor visual untuk menangkap fitur visual multi-tahap dari ViT, dan Q-former visual BLIP, yang menyelaraskan fitur gambar dengan ruang penyematan LLM melalui proyektor khusus bahasa.

Beberapa solusi berbasis GPT secara langsung menghasilkan kode yang dapat dieksekusi dari input visual. Voxposer menggunakan model besar multimodal untuk mengekstrak informasi objek geometris, menentukan koordinat 3D, dan menyematkannya ke dalam kode yang dapat dieksekusi untuk menghasilkan diagram fungsional dan kendala. ReKep meningkatkan adaptabilitas dengan mengintegrasikan DINOv2 untuk ekstraksi fitur tingkat patch, SAM untuk deteksi titik kunci, dan GPT-4o untuk perencanaan tugas multi-tahap dengan sub-tujuan dan kendala jalur. RobotGPT mengadopsi kerangka kerja petunjuk lima komponen terstruktur yang mencakup latar belakang, objek, lingkungan, tugas, dan contoh untuk memandu pembuatan kode berbasis GPT untuk kontrol robot yang presisi. Model besar non-GPT untuk persepsi yang terwujud meliputi PaLM-E, yang secara langsung mengintegrasikan input sensorik multimodal ke dalam representasi laten bersama yang diproses melalui perhatian diri Transformer, memungkinkan pengambilan keputusan sekuensial yang mendasar. Model seri RT, termasuk RT-1, RT-2, dan RT-H, secara progresif memajukan

pembelajaran visual-bahasa-aksi dengan mengkodekan input visual dengan ViT atau EfficientNet-B3 dan melatih pada dataset bahasa-visi skala besar untuk secara langsung menghasilkan aksi robot tingkat rendah.

Kontribusi penting lainnya termasuk RoboFlamingo, yang menyempurnakan Open-Flamingo untuk tugas manipulasi menggunakan data demonstrasi robot terbatas; Open-VLA, yang melengkapi Llama 2 dengan encoder visual DINOv2 dan SigLIP untuk meningkatkan penalaran spasial dan pemahaman semantik; dan RoboMamba, yang mengintegrasikan model Mamba dengan encoder visual untuk melatih penalaran visual-linguistik secara bersamaan untuk aplikasi robot.

Pembelajaran Kognitif Otonom

Pembelajaran kognitif otonom yang terwujud adalah domain interdisipliner yang mengambil inspirasi dari robotika, ilmu kognitif, kecerdasan buatan, ilmu saraf, dan psikologi perkembangan. Hal ini membahas tantangan mendasar dalam menciptakan sistem robot yang tidak hanya melakukan tindakan tetapi juga belajar dan beradaptasi melalui interaksi fisik dengan lingkungannya. Kognisi yang terwujud, gagasan bahwa kemampuan kognitif muncul bukan semata-mata dari komputasi simbolik internal tetapi dari interaksi antara tubuh dan dunia, merupakan inti dari bidang ini. Bab ini mengeksplorasi dasar-dasar teoritis, prinsip-prinsip arsitektur, mekanisme pembelajaran, implementasi praktis, dan arah masa depan pembelajaran kognitif otonom yang terwujud. Kita mulai dengan prinsip-prinsip kognisi yang terwujud, kemudian memeriksa robotika kognitif dan sistem pembelajaran perkembangan, dilanjutkan dengan kerangka kerja modern yang menggabungkan pembelajaran dan tindakan dalam sistem fisik, dan diakhiri dengan tren dan tantangan yang muncul.

Dasar-dasar Kognisi yang Terwujud. Kognisi yang terwujud menantang pandangan tradisional yang menempatkan kognisi semata-mata di dalam otak. Hal ini menekankan bahwa persepsi, memori, penalaran, dan bahasa didasarkan pada interaksi sensorimotor. Para pendukungnya berpendapat bahwa kognisi muncul dari interaksi antara tubuh agen, sistem sensorimotor, dan lingkungan. Beberapa sudut pandang teoretis mengilustrasikan pergeseran ini: (1) Psikologi Ekologis menyoroti bahwa persepsi muncul melalui interaksi langsung dengan peluang lingkungan. Agen tersebut merasakan peluang tindakan secara langsung, seperti mengenali bahwa kursi memungkinkan untuk duduk. (2) Simulasi Terwujud mengusulkan bahwa proses kognitif melibatkan simulasi pengalaman sensorimotor untuk memahami dan memprediksi hasil di lingkungan. (3) Enaktivisme membingkai kognisi sebagai aktif, situasional, dan muncul, menekankan bahwa kognisi muncul dari keterlibatan dinamis waktu nyata dengan dunia. Perspektif-perspektif ini menantang konseptualisasi abstrak simbolik dan tanpa tubuh tentang berpikir dan mendukung pandangan bahwa perilaku cerdas tidak dapat dipisahkan dari pengalaman tubuh.

Robotika Kognitif dan Pembelajaran Otonom. Robotika kognitif dibangun di atas kognisi yang terwujud dengan merancang sistem robot yang belajar dan bernalar melalui interaksi daripada mengandalkan aturan yang telah ditentukan sebelumnya. Prinsip-prinsip utama meliputi perwujudan, integrasi sensorimotor, pembelajaran daring, dan kemampuan beradaptasi. Pembelajaran seperti manusia bersifat bertahap, kumulatif, dan didorong oleh

rasa ingin tahu dan motivasi intrinsik. Robot yang dirancang dengan prinsip-prinsip ini menjelajahi lingkungannya secara otonom dan memilih tujuan mereka sendiri, secara bertahap membangun kemampuan dari waktu ke waktu. Pendekatan ini kontras dengan sistem pembelajaran mendalam yang biasanya membutuhkan kumpulan data berlabel yang besar dan tujuan statis. Robotika perkembangan mewujudkan visi ini dengan memodelkan lintasan pembelajaran yang mirip dengan yang diamati dalam perkembangan anak. Misalnya, mekanisme pembelajaran berlapis mereplikasi bagaimana seorang anak memperoleh keterampilan sensorimotor terlebih dahulu, kemudian membangunnya.

Perkembangan penting adalah integrasi model dunia dan pengkodean prediktif ke dalam kognisi robot. Melalui interaksi berkelanjutan, robot membangun model internal lingkungannya dan dirinya sendiri untuk memprediksi masukan sensorik di masa depan dan merencanakan tindakan yang sesuai. Mekanisme prediksi ini sejajar dengan bagaimana otak manusia mengantisipasi konsekuensi sensorik dari tindakan dan beradaptasi untuk meminimalkan kesalahan prediksi. Arsitektur kognitif yang canggih mereplikasi beberapa fungsi kognitif, persepsi, perencanaan, memori, dan bahasa, dalam kerangka kerja yang terwujud. Misalnya, arsitektur DAC-h3 memungkinkan robot humanoid untuk secara proaktif menjelajahi, belajar dari interaksinya, dan menceritakan pengalamannya.

Representasi dalam Pembelajaran Berbasis Wujud. Merepresentasikan pengetahuan dengan cara yang mendasar dan dapat digeneralisasikan tetap menjadi tantangan utama. Kerangka Kerja Kompleks Objek-Aksi (OAC) mengusulkan bahwa kognisi dapat dibangun di atas primitif representasional yang menggabungkan affordance (tindakan apa yang diizinkan oleh objek) dengan struktur perencanaan logis. OAC memfasilitasi pembelajaran hierarkis tindakan dengan mendasarkannya pada konteks persepsi dan tindakan. Neurorobotika menanamkan model yang terinspirasi secara neurobiologis ke dalam pengendali robot, memungkinkan mereka untuk belajar melalui mekanisme seperti plastisitas sinaptik, kontrol motorik hierarkis, dan sistem memori. Sistem tersebut mensimulasikan fungsi seperti otak seperti generator pola pusat untuk lokomosi, peta hipokampus untuk navigasi, dan modul persepsi sensorik. Sistem ini menggambarkan bagaimana kognisi berbasis wujud mendapat manfaat dari kerangka kerja yang terinspirasi secara biologis.

Pembelajaran Berbasis Wujud dalam Praktik. Karya terbaru telah secara meyakinkan menunjukkan pembelajaran kognitif berbasis wujud pada robot. Sistem ELLMER memanfaatkan GPT-4 dan infrastruktur generasi yang diperkuat dengan pengambilan untuk menghasilkan rencana yang sadar konteks yang menggabungkan umpan balik selama pelaksanaan tugas. Robot yang dilengkapi dengan ELLMER berhasil menyelesaikan tugas-tugas jangka panjang seperti pembuatan kopi dan dekorasi piring dengan beradaptasi berdasarkan umpan balik gaya dan penglihatan. Dalam robotika bedah, pembelajaran berbasis visi memungkinkan otomatisasi sub-tugas umum di berbagai prosedur. Model yang dilatih dalam paradigma pembelajaran yang terwujud telah menunjukkan kemampuan untuk melakukan generalisasi di luar tugas-tugas spesifik yang dilatih. Sistem AI berwujud semakin banyak digunakan dalam konteks dunia nyata. Misalnya, robot di Tiongkok termasuk humanoid dan drone pengiriman beroperasi secara otonom di ruang publik, belajar dari masukan sensorik

dan menyesuaikan perilaku sesuai dengan itu. Integrasi AI berwujud ke dalam kehidupan sehari-hari menggarisbawahi relevansi dan dampak sosial dari pembelajaran kognitif yang terwujud.

Pembelajaran kognitif otonom yang terwujud menghadirkan kerangka kerja holistik dan ampuh untuk membangun sistem cerdas yang belajar melalui pengalaman fisik dan kontekstual. Kombinasi perwujudan, pembelajaran perkembangan yang termotivasi secara intrinsik, pemodelan prediktif, representasi adaptif, dan interaksi sosial menghasilkan kecerdasan yang tangguh dan fleksibel pada robot. Seiring kemajuan kita dalam model komputasi, integrasi sensorimotor, pemahaman bahasa, dan interaksi manusia-robot, pembelajaran kognitif otonom yang terwujud akan tetap menjadi paradigma vital yang memandu pengembangan mesin cerdas di masa depan.

Interaksi Dinamis Otonom

Gagasan tentang agen robotik yang berwujud muncul pada akhir tahun 1980-an, berakar pada pemahaman bahwa perilaku cerdas muncul dari perwujudan fisik suatu sistem dan interaksi sensorimotornya yang berkelanjutan dengan lingkungan sekitarnya. Perspektif ini menginspirasi arsitektur neural ujung-ke-ujung awal yang secara langsung memetakan input sensorik ke output motorik, memungkinkan kontrol otonom tanpa penalaran simbolik eksplisit. Pengenalan model bahasa besar (LLM) telah secara substansial memperluas kemampuan fungsional agen otonom, menggesernya dari sistem berbasis aturan yang ruang lingkupnya sempit menuju entitas fleksibel dengan kemampuan untuk menggeneralisasi, bernalar dalam konteks, dan beradaptasi dengan situasi baru. Sebelum era ChatGPT, Code as Policies (CaP) mempelopori konsep program model bahasa yang dapat dieksekusi robot (LMP), memungkinkan kontrol reaktif dan berbasis visi dengan menghasilkan kode yang secara tidak langsung memanggil alat robotik.

Rilis ChatGPT OpenAI pada November 2022 menandai titik balik dalam riset agen berwujud, mempercepat adopsi LLM sebagai komponen inti kerangka kerja kontrol robot. Meskipun Robot Operating System (ROS) telah lama berfungsi sebagai middleware de facto untuk integrasi robot, prototipe LLM-robotika awal terutama menargetkan platform yang kompatibel dengan ROS. Pada tahun 2023, implementasi awal berfokus pada pembentukan integrasi protokol dan saluran komunikasi dua arah antara LLM dan ROS. Contohnya termasuk ekstensi bahasa alami untuk ROS 2 seperti *ros2ai*, yang melengkapi antarmuka baris perintah ROS 2 dengan kemampuan LLM, dan ROScribe, yang menghasilkan basis kode ROS dari deskripsi tingkat tinggi arsitektur sistem yang diinginkan. Kerangka kerja ROS-LLM menyediakan templat yang dapat dikonfigurasi untuk mendefinisikan API robot, secara otomatis menangani konstruksi prompt dan panggilan API untuk memfasilitasi interaksi bahasa alami. Bersamaan dengan itu, ChatGPT Microsoft untuk Robotika memperkenalkan desain modular untuk pemrograman robot berbasis LLM, mendefinisikan API tingkat tinggi per robot dan memanfaatkan rekayasa cepat dengan pustaka fungsi yang telah ditentukan sebelumnya untuk memungkinkan kontrol bahasa alami di berbagai tugas dan lingkungan. Sebagian besar sistem ini mengadopsi antarmuka bergaya chatbot, menafsirkan input pengguna bentuk bebas dan menghasilkan perintah robot yang sesuai.

Kemajuan pada tahun 2024 memperkenalkan antarmuka LLM-robotika khusus dan sistem visi-bahasa-aksi (VLA) yang mencapai tingkat generalisasi dan penalaran semantik yang belum pernah terjadi sebelumnya. Sistem ini dapat memproses instruksi bahasa terbuka, mendasarkannya pada persepsi visual, dan mengeksekusi tindakan yang relevan. RT-2 mendemonstrasikan strategi penyempurnaan bersama yang menggabungkan model yang telah dilatih sebelumnya dengan data lintasan robot, memungkinkan persepsi, penalaran, dan kontrol terintegrasi. Sistem π *0* lebih lanjut menyatukan persepsi, penalaran, dan pembangkitan gerakan melalui kontrol kontinu waktu nyata berdasarkan kebijakan difusi yang sesuai dengan aliran. Secara paralel, kerangka kerja middleware baru muncul untuk menjembatani penalaran LLM dan tumpukan perangkat lunak robot tradisional. ROSA (Robot Operating System Agent) mengimplementasikan lapisan kontrol berbasis LLM menggunakan LangChain dan paradigma ReAct, mengabstraksi operasi ROS ke dalam alat berbasis Python untuk tugas-tugas seperti kueri sensor dan kontrol aktuator, sambil menyematkan mekanisme keselamatan seperti validasi parameter dan persetujuan manusia. RAI (Robotic AI Agent) memperkenalkan arsitektur multi-agen terdistribusi, di mana agen LLM khusus, termasuk modul persepsi, perencanaan, kontrol, dan keselamatan, berkolaborasi secara waktu nyata. BUMBLE (BUilding-wide MoBiLE Manipulation) menerapkan kerangka kerja berbasis VLM terpadu untuk manipulasi seluler skala besar, mengintegrasikan persepsi dunia terbuka, sistem memori, dan beragam keterampilan motorik untuk beroperasi pada skala spasial dan temporal yang diperluas.

Garis pengembangan lain berfokus pada server Model Context Protocol (MCP), yang memungkinkan asisten AI untuk berintegrasi secara mulus dengan alat robotika. Pustaka ROS-MCP, misalnya, menghubungkan asisten seperti Claude ke sistem ROS 2, memungkinkan manajemen sumber daya, kontrol simulator, dan debugging melalui antarmuka bahasa alami. Pendekatan berbasis plugin ini kontras dengan kerangka kerja agen monolitik, menawarkan ekstensibilitas modular yang lebih besar. Baru-baru ini, OM1 OpenMind memperkenalkan runtime AI yang tidak bergantung pada perangkat keras yang dibangun di atas protokol koordinasi terdesentralisasi, mendukung manajemen identitas yang aman dan interoperabilitas di antara platform robot heterogen. Meskipun mengadopsi loop persepsi-aksi yang mirip dengan sistem VLA, OM1 menekankan koordinasi terdistribusi daripada orkestrasi terpusat. Selama tiga tahun terakhir, perkembangan ini telah menetapkan tren yang jelas menuju pengintegrasian lapisan "agen AI" ke dalam sistem robot, yang menafsirkan niat manusia, bernalar atas input multimodal, dan secara otonom mengelola kemampuan perangkat lunak robot.

Perwujudan Perangkat Keras Otonom

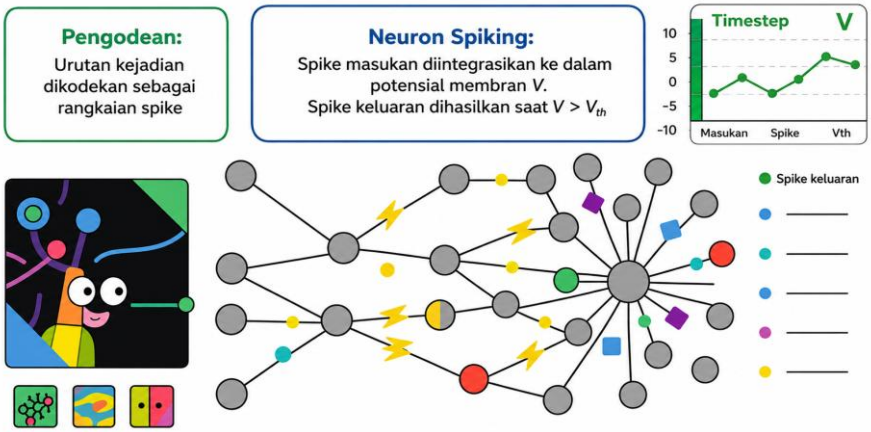
Konsep Perwujudan Perangkat Keras Otonom mengacu pada integrasi sistem kontrol cerdas dengan tubuh robot fisik, yang memungkinkan agen untuk merasakan, bernalar, dan bertindak di dunia nyata tanpa intervensi manusia secara langsung. Tidak seperti model AI berbasis perangkat lunak murni, perwujudan otonom harus mengatasi kompleksitas koordinasi sensorimotor, keandalan perangkat keras, dan adaptasi lingkungan secara waktu nyata. Paradigma ini terletak di persimpangan robotika, kecerdasan buatan, mekatronika, dan

teori kontrol, dan memainkan peran penting dalam berbagai bidang mulai dari otomatisasi industri dan kendaraan otonom hingga robotika layanan dan riset AI berwujud. Gagasan perwujudan menekankan bahwa kecerdasan bukan semata-mata fungsi komputasi abstrak tetapi muncul dari keterkaitan erat antara tubuh fisik, persepsi sensorik, dan interaksi lingkungan. Dalam sistem perangkat keras otonom, faktor bentuk perangkat keras, kemampuan aktuasi, dan penempatan sensor bukan hanya kendala teknik, tetapi juga membentuk sifat dasar kognisi dan kinerja agen.

Komputasi Neuromorfik

Permintaan akan sistem otonom yang cerdas, adaptif, dan hemat energi di berbagai bidang seperti robotika, agen bergerak (misalnya, UAV), dan kendaraan swakemudi telah meningkat secara signifikan dalam beberapa tahun terakhir, yang memicu minat yang semakin besar pada komputasi neuromorfik. Mengambil inspirasi dari prinsip-prinsip komputasi sistem saraf biologis, pendekatan neuromorfik memberikan jalur yang menjanjikan untuk meningkatkan persepsi, pengambilan keputusan, dan responsivitas platform otonom. Dalam aplikasi praktis, sistem seperti kendaraan otonom, drone, dan perangkat yang dapat dikenakan harus membuat keputusan secara waktu nyata di bawah kendala ketat pada konsumsi energi dan sumber daya komputasi. Arsitektur komputasi konvensional sering menghadapi keterbatasan dalam memenuhi persyaratan ini karena tuntutan daya yang tinggi dan latensi pemrosesan. Komputasi neuromorfik mengatasi tantangan ini melalui penggunaan Jaringan Saraf Spiking (SNN), yang memproses informasi sebagai peristiwa diskrit, memungkinkan inferensi daya rendah dan latensi rendah. Paradigma berbasis peristiwa ini secara alami cocok untuk memproses data yang jarang secara temporal dan terintegrasi secara efektif dengan modalitas penginderaan berbasis peristiwa. Terletak di persimpangan robotika, ilmu saraf, pembelajaran mesin, dan desain perangkat keras, komputasi neuromorfik menawarkan kerangka kerja multidisiplin untuk memajukan kemampuan sistem otonom generasi berikutnya.

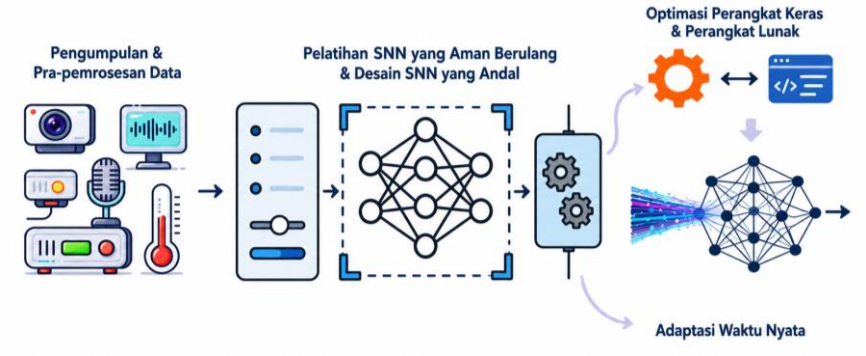
Jaringan Saraf Spiking. Jaringan Saraf Spiking (SNN), yang diakui sebagai generasi ketiga model jaringan saraf, dirancang untuk menangkap dinamika neuron biologis dengan lebih akurat. Dalam SNN, informasi ditransmisikan melalui spike diskrit daripada nilai aktivasi kontinu. Setiap neuron mengakumulasi spike yang masuk ke dalam potensial membran dan memancarkan spike ketika ambang batas yang ditentukan terlampaui, seperti yang diilustrasikan pada Gambar 2.6. Komputasi berbasis spike ini secara alami mengkodekan ketergantungan temporal, memungkinkan pengenalan pola yang bervariasi waktu secara efisien. Karakteristik tersebut membuat SNN sangat cocok untuk aplikasi yang membutuhkan respons latensi rendah dan sensitivitas temporal, termasuk deteksi gerakan dan pengambilan keputusan waktu nyata dalam sistem otonom.



Gambar 2.6 Cara kerja dasar jaringan saraf spiking

Perangkat Keras Neuromorfik. Penerapan SNN yang efisien bergantung pada perangkat keras neuromorfik, yang mengimplementasikan komputasi jaringan dengan meniru konektivitas struktural dan dinamika fungsional sistem saraf biologis. Arsitektur seperti IBM TrueNorth, Intel Loihi, SpiNNaker, dan Brain-ScaleS, serta solusi komersial seperti Brainchip Akida dan Synsense DYNAP-SE, mencontohkan pendekatan ini. Platform ini dioptimalkan untuk operasi throughput tinggi dan daya rendah, sehingga memungkinkan untuk menerapkan model SNN skala besar di lingkungan yang terbatas sumber daya. Dengan mengintegrasikan efisiensi algoritma dengan desain perangkat keras khusus, komputasi neuromorfik menciptakan fondasi yang terukur untuk persepsi dan kontrol otonom yang hemat energi.

Seperti yang ditunjukkan pada Gambar 2.7, alur kerja yang diusulkan dimulai dengan akuisisi data mentah melalui sensor tradisional atau berbasis peristiwa, diikuti oleh tahap pra-pemrosesan yang mencakup diskretisasi dan pengkodean peristiwa. Menghasilkan urutan peristiwa terstruktur temporal yang cocok untuk pemrosesan neural spiking. Selanjutnya, model SNN yang andal dirancang, menggabungkan penyetelan hiperparameter dan toleransi kesalahan untuk ketahanan terhadap variabilitas input dan perilaku sistem. Setelah itu, pipeline pelatihan yang aman menggabungkan pelatihan adversarial dengan mekanisme enkripsi untuk menghasilkan SNN yang tangguh. Selanjutnya, optimasi tingkat perangkat keras dan perangkat lunak, seperti pemangkasan, kuantisasi, dan NAS yang sadar perangkat keras, diterapkan, dipandu oleh batasan desain chip neuromorfik. Langkah terakhir memperkenalkan kemampuan adaptasi waktu nyata seperti pembelajaran berkelanjutan dan penyaringan noise, sehingga model dapat tetap efisien dalam kondisi dinamis. Tahapan-tahapan ini menghasilkan sistem SNN yang aman, efisien, dan andal yang mendukung kecerdasan terwujud dalam sistem otonom.



Gambar 2.7 Alur kerja jaringan saraf spiking

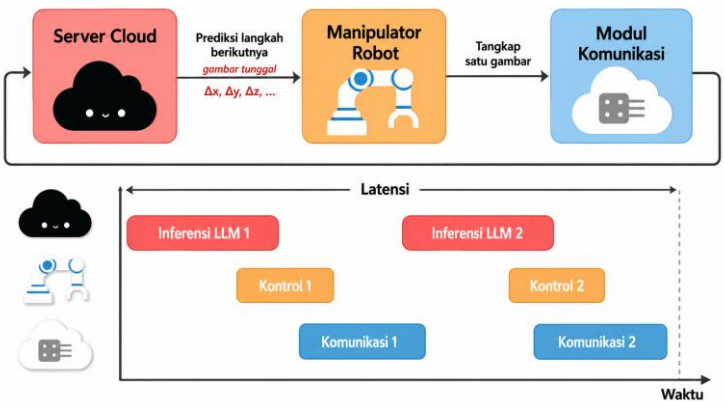
Berbasis LLM

Robot tradisional biasanya mengandalkan algoritma berbasis aturan untuk pengambilan keputusan dan perencanaan tugas, yang membatasi penerapannya pada skenario yang relatif sederhana dan telah ditentukan sebelumnya. Kemajuan pesat Model Bahasa Besar (LLM) telah menginspirasi pendekatan baru yang bertujuan untuk memberikan robot kemampuan penalaran tingkat lanjut dan perencanaan jangka panjang. Kemajuan ini telah mendorong pengembangan aplikasi yang mengintegrasikan LLM ke dalam kontrol robot, menghasilkan peningkatan kinerja yang signifikan, khususnya dalam meningkatkan tingkat keberhasilan robot dalam melakukan tugas-tugas kompleks di lingkungan dinamis. Sistem AI yang Terwujud (EAI) membentuk kategori sistem cerdas yang memanfaatkan kemampuan penalaran LLM untuk memandu robot dalam melakukan tugas-tugas dunia nyata yang kompleks seperti pekerjaan rumah tangga dan manufaktur industri, sehingga mengurangi kebutuhan intervensi manusia. Sistem ini umumnya terdiri dari dua komponen utama: server berkinerja tinggi dengan GPU untuk inferensi LLM dan platform robot yang mengeksekusi perintah dan berinteraksi dengan lingkungan fisik. Sebagian besar sistem EAI mengadopsi LLM multimodal sebagai modul perencanaan. Model ini mengintegrasikan instruksi bahasa alami, misalnya “letakkan cangkir biru di atas meja dan bawakan saya yang merah,” dengan aliran data sensor konvensional dalam alur kerja robot, termasuk video kontinu, sinyal IMU, dan awan titik.

LLM memproses input multimodal ini dan menghasilkan urutan tindakan untuk dieksekusi oleh robot, berdasarkan instruksi yang diberikan dan pengamatan sensorik terbaru. Kemajuan terbaru dalam EAI telah menunjukkan potensi signifikan untuk menggantikan manusia dalam berbagai tugas. Misalnya, Robotic Transformer milik Google telah mencapai tingkat keberhasilan melebihi 75,0% dalam tugas-tugas rumah tangga seperti “mengambil benda,” “membuka laci,” dan “menempatkan benda ke lokasi yang ditentukan.” RoboFlamingo, kerangka kerja EAI yang baru diperkenalkan, telah lebih meningkatkan tingkat keberhasilan tugas tunggal menjadi lebih dari 89,5%. Pencapaian ini menunjukkan bahwa integrasi LLM ke dalam sistem robot untuk interaksi fisik dunia nyata merupakan langkah menjanjikan menuju realisasi Kecerdasan Buatan Umum (AGI).

Perbedaan utama antara menerapkan LLM untuk pembuatan teks atau gambar dan mengintegrasikannya sebagai modul pengambilan keputusan dan perencanaan dalam sistem

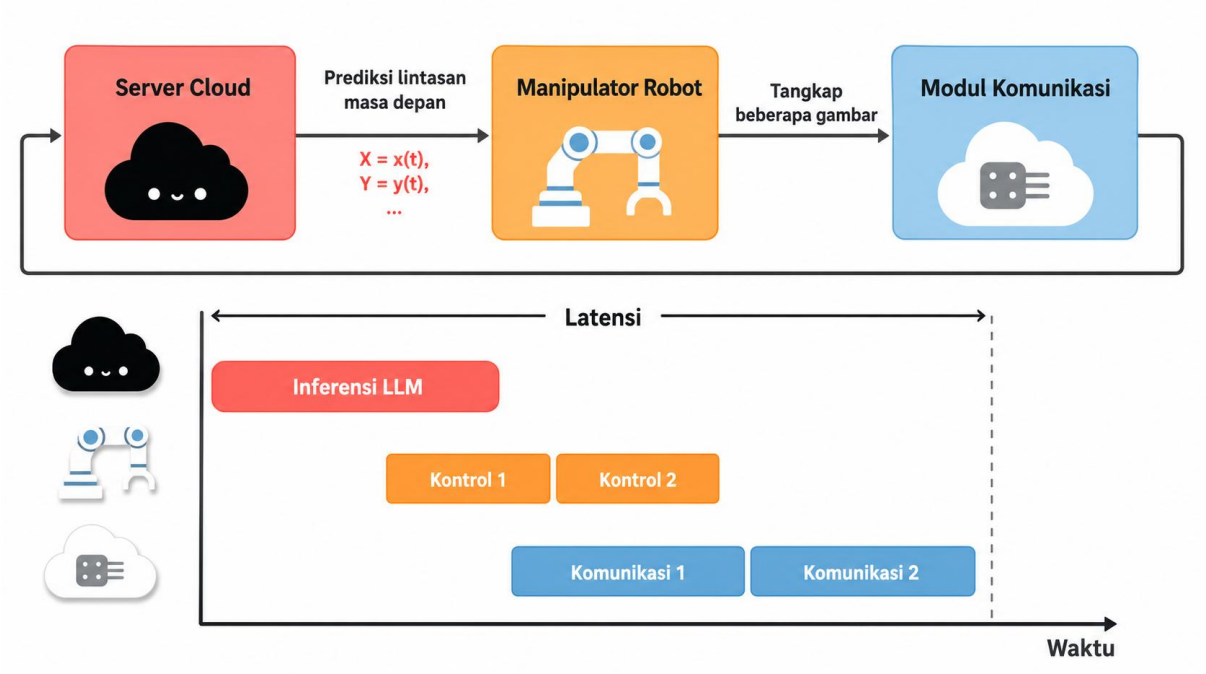
robot terletak pada kendala waktu nyata yang ketat yang harus dipenuhi robot. Tanpa kinerja waktu nyata yang terjamin, penerapan sistem EAI tetap terbatas pada eksplorasi teoretis daripada penerapan praktis. Namun, memenuhi kendala waktu nyata tetap menjadi tantangan signifikan bagi sistem EAI saat ini. Akar penyebabnya terletak pada paradigma eksekusinya. Sistem EAI yang ada sebagian besar mengadopsi model eksekusi sekuensial, di mana input video diproses, dan tindakan robot dihasilkan berdasarkan bingkai demi bingkai. Secara khusus, setelah inisialisasi, robot dimulai dengan urutan video yang berisi N bingkai dan instruksi bahasa i. LLM memprediksi tuple aksi $(\Delta x, \Delta y, \Delta z, \ldots)$ berdasarkan tuple input saat ini $(Frame_{t-N}, Framet - N + 1, \ldots, Frame_t, i)$, di mana Δ mewakili gerakan yang diusulkan dan $Frame_t$ menunjukkan gambar terbaru dalam urutan tersebut. Setelah mengeksekusi aksi yang diprediksi, robot menangkap frame berikutnya $Frame_{t+1}$ pada posisi yang diperbarui. Inferensi LLM berikutnya kemudian memproses tuple input yang diperbarui $(Frame_{t-N+1}, Frame_{t-N+2}, \ldots, Frame_{t+1}, i)$. Model eksekusi sekuensial ini pada dasarnya memakan waktu karena dua alasan utama. Seperti yang ditunjukkan pada Gambar 2.8, pertama, ketergantungan setiap tahap pada tahap sebelumnya menyebabkan latensi kumulatif yang substansial. Karena sebagian besar robot bergantung pada server jarak jauh berkinerja tinggi untuk inferensi LLM, total penundaan untuk setiap frame dihasilkan dari efek gabungan latensi inferensi LLM, latensi eksekusi aksi robot, dan latensi transmisi data, yang dapat berjumlah beberapa ratus milidetik. Kedua, karena ketiga tahap tersebut harus diselesaikan untuk setiap frame, responsivitas sistem semakin terbatas, sehingga sulit untuk mencapai operasi waktu nyata yang sebenarnya.



Gambar 2.8 Alur eksekusi diskrit dari sistem AI berwujud

Untuk mengurangi latensi ujung-ke-ujung dari sistem AI yang ada, CORKI mendesain ulang pipeline eksekusi. Seperti yang diilustrasikan pada Gambar 2.9, peningkatan pertama terjadi pada tingkat algoritma. CORKI menyimpang dari paradigma konvensional yang memprediksi pergerakan robot secara diskrit untuk frame berikutnya dan sebagai gantinya mengusulkan kerangka kerja AI baru yang mampu memprediksi lintasan robot dalam waktu dekat. Tidak seperti metode yang hanya memprediksi langkah selanjutnya, kerangka kerja ini secara akurat menghasilkan urutan aksi yang mencakup beberapa langkah ke depan. Dengan demikian, hal ini secara substansial menurunkan frekuensi inferensi LLM, sehingga mengurangi

latensi dan konsumsi energi. Peningkatan kedua menargetkan proses kontrol. CORKI mengembangkan akselerator khusus yang mengubah lintasan yang diprediksi menjadi sinyal kontrol waktu nyata yang kontinu. Akselerator ini secara khusus dioptimalkan untuk kontrol torsi komputasi ruang tugas dan menggabungkan arsitektur aliran data yang disesuaikan, sirkuit khusus, dan desain buffer on-chip khusus.



Gambar 2.9 Alur eksekusi berkelanjutan dari sistem AI yang diwujudkan

Terakhir, pada tingkat sistem, CORKI mengoptimalkan transmisi frame yang baru ditangkap dengan tumpang tindih proses komunikasi dengan proses eksekusi robot. Desain ini secara efektif menyembunyikan latensi komunikasi di dalam latensi eksekusi, sehingga meminimalkan penundaan ujung-ke-ujung secara keseluruhan. Perwujudan Perangkat Keras Otonom mewakili konvergensi desain mekanik, penginderaan, komputasi, dan kontrol berbasis AI menjadi satu kesatuan yang kohesif yang mampu beroperasi secara independen di lingkungan yang kompleks. Badan perangkat keras membentuk kemampuan persepsi dan tindakan agen, sementara kecerdasan yang tertanam memungkinkan adaptasi, ketahanan, dan perilaku yang bertujuan. Seiring dengan matangnya teknologi pendukung, kita dapat mengharapkan perwujudan yang lebih adaptif, hemat energi, dan sadar sosial, mempercepat integrasi sistem otonom ke dalam setiap aspek kehidupan manusia.

BAB 3
AI BERWUJUD KAWANAN OTONOM

3.1 PENDAHULUAN

AI Berwujud Kawanan Otonom (*Autonomous Swarm Embodied Artificial Intelligence*/Autonomous SEAI) merepresentasikan evolusi signifikan dari sistem agen tunggal menuju sistem multiagen dengan memanfaatkan kecerdasan kawanan untuk mengatasi keterbatasan persepsi lingkungan, pengambilan keputusan terwujud, dan ketahanan operasional pada AI terwujud saat ini. Bab ini meninjau AI berwujud kawanan dari fondasi hingga perkembangan mutakhir, termasuk teknologi utama, komponen utama, ketahanan operasional pada pengaturan sim-to-real, simulator, dataset, dan metrik evaluasi. Kemajuan terbaru dalam AI berwujud otonom telah sangat memajukan bidang ini. Bab ini menyoroti pendekatan kolaborasi otonom, komunikasi otonom, dan optimasi ketahanan otonom dalam Autonomous SEAI melalui tiga paradigma integratif: pengorganisasian diri, negosiasi diri, dan refleksi diri. Autonomous SEAI akan menjadi arah riset penting pada masa mendatang karena mampu mempertahankan resiliensi, adaptabilitas, dan perbaikan diri berkelanjutan dalam lingkungan dinamis.

Dekade terakhir telah menyaksikan lonjakan riset dan aplikasi Kecerdasan Buatan yang Terwujud (*Embodied Artificial Intelligence*/EAI) dalam berbagai skenario dunia nyata, seperti robot industri untuk penyortiran logistik, pengelasan otomotif, dan pengecoran beton; robot rumah tangga untuk menyedot debu, membersihkan, dan prostetik; dan robot khusus untuk patroli perbatasan, eksplorasi bawah air, dan konservasi biologi. Sebagian besar metode EAI bertujuan untuk memungkinkan agen tunggal untuk secara sempurna melakukan tugas spesifik dalam skenario spesifik melalui persepsi, penalaran, pengambilan keputusan, dan interaksi tindakan. Namun, kinerja mereka sering terganggu oleh oklusi, degradasi atau kegagalan sensor, dan keterbatasan mode diri, dan mereka umumnya kurang memiliki kemampuan untuk persepsi jarak jauh dan pengambilan keputusan jangka panjang. Selain itu, pilihan model optimal sering bervariasi tergantung pada skenario spesifik, mode ego, dan tugas, yang membutuhkan keahlian substansial untuk optimasi adaptif. Kompleksitas ini menimbulkan hambatan signifikan bagi pengguna baru dan membatasi penerapan dan kelayakan metode EAI modern di lingkungan yang dinamis dan terbuka.

AI Berwujud Kawanan Otonom (*Autonomous Swarm Embodied Artificial Intelligence*/Autonomous SEAI) menandai evolusi signifikan dalam AI berwujud kawanan, bergeser dari operasi agen tunggal ke kerangka kerja kecerdasan kawanan. Dalam kerangka kerja ini, kemampuan tingkat sistem yang canggih muncul melalui koordinasi terdesentralisasi di antara banyak agen fisik. Dengan mengintegrasikan penginderaan, komunikasi, dan perilaku adaptif yang terdistribusi, Autonomous SEAI secara efektif mengatasi banyak kendala yang melekat pada sistem agen tunggal, khususnya di lingkungan yang dinamis dan terbuka. Pendekatan kolaboratif ini mendukung peningkatan kesadaran situasional, peningkatan

kualitas pengambilan keputusan, dan ketahanan yang lebih besar terhadap ketidakpastian dan gangguan dunia nyata.

Mengambil inspirasi dari kolektif biologis seperti serangga sosial dan burung yang berkelompok, SEAI memformalkan bagaimana interaksi lokal menghasilkan kecerdasan kawanan yang muncul. Yang sentral dalam SEAI adalah peran perwujudan fisik dan interaksi agen-lingkungan, yang bersama-sama memfasilitasi koordinasi yang koheren dan sadar konteks. Kerangka kerja ini menunjukkan sifat-sifat penting termasuk respons adaptif, stabilitas operasional, dan kinerja yang dapat diskalakan, sering kali menunjukkan peningkatan nonlinier seiring bertambahnya jumlah agen. Kapasitas kolektif tersebut memungkinkan sistem SEAI untuk mengatasi tantangan yang tidak mungkin dilakukan oleh agen individual, terutama dalam skenario skala besar atau terdistribusi secara spasial.

Teknologi utama SEAI meliputi mekanisme untuk pengambilan keputusan kolektif, regulasi terdistribusi, dan sinkronisasi yang terorganisasi sendiri. Algoritma seperti Optimasi Kawanan Partikel (PSO) menerjemahkan perilaku swarm alami ke dalam strategi komputasi untuk memecahkan tugas optimasi yang kompleks. Metode ini memungkinkan sistem terdesentralisasi untuk mencapai tujuan global melalui aturan lokal dan umpan balik iteratif, mempertahankan fungsionalitas di tengah variabilitas lingkungan dan gangguan internal. Komponen utama SEAI menggabungkan kerangka kerja kolaborasi multi-tahap. Kolaborasi awal memanfaatkan fusi data mentah untuk memaksimalkan integritas persepsi dengan mengorbankan bandwidth. Kolaborasi menengah memanfaatkan pertukaran tingkat fitur untuk menyeimbangkan detail dan efisiensi, sementara kolaborasi akhir beroperasi pada output tingkat keputusan untuk meminimalkan overhead komunikasi. Modul khusus untuk komunikasi dinamis, integrasi fitur, dan kustomisasi kerugian lebih lanjut menyempurnakan persepsi kolaboratif dan memungkinkan kinerja yang tangguh.

Ketahanan operasional SEAI merupakan perhatian vital untuk aplikasi dunia nyata. Tantangan utama meliputi ketidakakuratan pose, konektivitas yang terputus-putus, varians model antar agen, dan ancaman keamanan siber. Solusi kontemporer menggunakan jaringan penyempurnaan pose, protokol komunikasi yang toleran terhadap latensi, penyelarasan fitur lintas domain, dan berbagi data kriptografi untuk memperkuat daya tahan dan keamanan sistem secara keseluruhan. SEAI otonom mencakup tiga modul inti: kolaborasi otonom, komunikasi otonom, dan optimasi ketahanan otonom, yang terwujud melalui tiga paradigma integratif: pengorganisasian diri, negosiasi, dan refleksi. Pengorganisasian diri memfasilitasi koordinasi yang muncul melalui interaksi lokal dan lingkaran umpan balik. Negosiasi menyediakan protokol eksplisit untuk resolusi konflik dan alokasi tugas. Refleksi memungkinkan pembelajaran berkelanjutan dari interaksi historis, mendorong adaptasi strategis dan peningkatan sistem jangka panjang.

Kemajuan dalam SEAI Otonom sangat bergantung pada alat simulasi yang canggih, kumpulan data yang kaya akan anotasi, dan metrik kinerja yang komprehensif. Simulator pemetaan dunia nyata seperti U2USim, bersama dengan kumpulan data multimodal skala besar seperti U2UData, menawarkan platform penting untuk pengembangan dan validasi algoritma. Sumber daya ini membantu mempersempit kesenjangan realitas, memastikan

bahwa model kecerdasan kawanan dapat ditransfer secara efektif dari simulasi ke penerapan dunia nyata. Kesimpulannya, SEAI Otonom mensintesis persepsi kolaboratif, kontrol terdistribusi, dan pengoptimalan diri berkelanjutan untuk menciptakan sistem multi-agen yang tangguh dan terukur. Dengan memanfaatkan kecerdasan kawanan, SEAI Otonom melampaui keterbatasan sistem agen tunggal tradisional dan menetapkan paradigma baru untuk operasi cerdas yang terwujud dalam lingkungan yang dinamis dan terbuka.

3.2 KETERBATASAN UTAMA AGEN TUNGGAL YANG TERWUJUD

AI yang Terwujud mengacu pada sistem siklus tertutup yang terdiri dari persepsi, tindakan, dan umpan balik, yang diwujudkan melalui interaksi fisik antara agen dan lingkungannya. AI yang tidak berwujud secara tradisional sering memisahkan kognisi dari tubuh, sedangkan AI yang berwujud menekankan bahwa esensi kecerdasan terletak pada adaptasi dan penyesuaian berkelanjutan dalam lingkungan dunia nyata yang dinamis. Pendekatan ini sangat mirip dengan pola perilaku alami manusia dan organisme lainnya. Dalam buku ini, AI yang berwujud secara khusus berfokus pada konsep agen tunggal. Agen tunggal mengacu pada entitas otonom yang mampu secara independen merasakan lingkungan, membuat keputusan, dan mengeksekusi tindakan. Model AI perwujudan individu ini telah menjadi aspek fundamental dari riset AI yang berwujud. Dengan berfokus pada perilaku agen tunggal, pendekatan ini memberikan kerangka kerja yang jelas dan intuitif untuk memahami bagaimana kecerdasan beroperasi melalui interaksi fisik dengan lingkungan.

Namun, meskipun agen individu dapat menunjukkan tingkat perilaku cerdas tertentu di berbagai lingkungan, keterbatasan dan tantangan mereka menjadi semakin jelas ketika menghadapi tugas-tugas yang kompleks. Bab ini membahas kekurangan-kekurangan ini dan memberikan analisis komprehensif dari berbagai perspektif: Keterbatasan Persepsi dan Adaptasi Lingkungan, Keterbatasan Kapasitas Pembelajaran dan Penalaran yang Berwujud, dan Keterbatasan Ketahanan Operasional.

Keterbatasan Persepsi Lingkungan dan Kapasitas Adaptasi

Agen individu beroperasi dengan rasionalitas terbatas, hanya mempersepsikan informasi parsial dan bising dari lingkungannya. Persepsi yang terbatas ini membatasi pemahaman mereka pada data lokal atau abstrak, sehingga mencegah pemahaman penuh terhadap sistem yang kompleks. Terlepas dari keterbatasan ini, agen menunjukkan kapasitas adaptif dengan menerapkan aturan sederhana, belajar dari umpan balik, dan menyesuaikan strategi berdasarkan pengalaman subjektif mereka. Mereka membentuk model internal untuk memprediksi hasil, memprioritugasan stimulus yang relevan, dan secara bertahap meningkatkan keputusan dalam konteks sempit mereka. Meskipun adaptasi sering kali efektif secara lokal, hal itu dapat menyebabkan hasil global yang suboptimal karena pengetahuan yang tidak lengkap. Keseimbangan antara persepsi terbatas dan perilaku adaptif ini sangat penting bagi fungsionalitas mereka dalam sistem multi-agen dan dunia nyata.

Kemampuan Persepsi Terbatas

AI yang Terwujud membutuhkan sistem persepsi untuk memahami lingkungan fisik dan berinteraksi dengannya. Sistem persepsi ini sebagian besar bergantung pada banyak sensor

untuk mengumpulkan informasi multimodal, memungkinkan agen untuk merasakan, memproses, dan memahami lingkungan. Dalam AI yang Terwujud secara individual, sensor tersebut menghadapi keterbatasan seperti jumlah yang kecil dan cakupan yang terbatas. Keterbatasan ini sangat bermasalah dalam skenario yang membutuhkan kesadaran situasional yang tinggi dan dinamika waktu nyata.

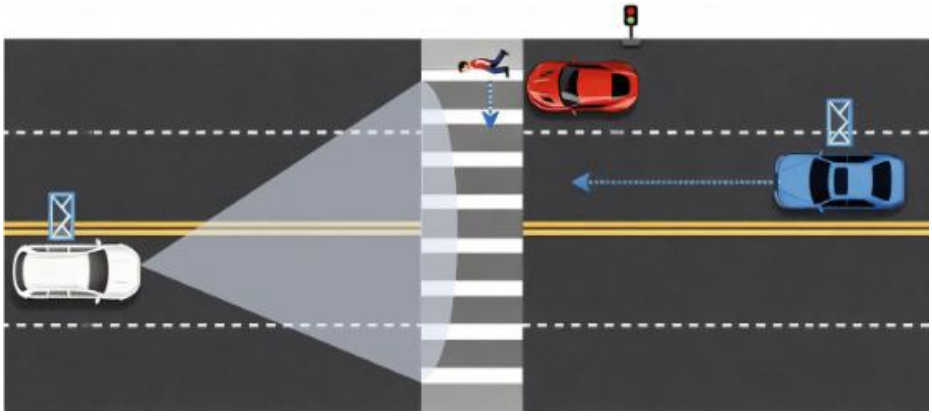
Akurasi Sensor dan Dampaknya. Salah satu tantangan utama yang dihadapi oleh EAI adalah akurasi sistem sensornya. Akurasi sensor mengacu pada kemampuan agen untuk mendeteksi dan mengukur rangsangan lingkungan secara akurat. Misalnya, sensor kamera mungkin mengalami kesulitan dalam kondisi cahaya rendah atau saat mendeteksi objek dari jarak jauh; sensor taktil mungkin kesulitan untuk menilai tekanan atau tekstur permukaan secara akurat. Jika sensor kurang akurat, agen mungkin menerima data yang tidak lengkap atau salah, yang menyebabkan kesalahpahaman tentang lingkungannya. Masalah ini menjadi sangat jelas dalam tugas-tugas dengan jenis sensor yang tidak memadai. Tugas yang berbeda biasanya membutuhkan beberapa jenis masukan sensor untuk sepenuhnya memahami lingkungan dan membuat keputusan yang tepat. Misalnya, agen yang hanya mengandalkan sensor visual mungkin tidak dapat mendeteksi faktor-faktor kunci seperti suhu, kelembaban, atau komposisi gas, yang sangat penting dalam misi industri atau penyelamatan tertentu.

Untuk memenuhi tuntutan tugas-tugas kompleks ini, mungkin perlu untuk mengintegrasikan jenis sensor lain, seperti pencitraan termal, deteksi gas, atau sensor akustik, untuk meningkatkan persepsi lingkungan agen secara keseluruhan. Selain itu, kegagalan sensor adalah tantangan utama lainnya. Seiring waktu atau dalam lingkungan ekstrem, sensor dapat mengalami degradasi atau malfungsi, yang mengakibatkan kehilangan atau distorsi data. Sebagai contoh, sensor kamera dapat gagal berfungsi di bawah cahaya yang kuat atau suhu ekstrem, sementara sensor taktil dapat kehilangan sensitivitas karena kerusakan mekanis atau keausan. Ketika sensor utama gagal, agen mungkin tidak dapat mengumpulkan data lingkungan yang akurat, sehingga mengganggu fungsinya, menyebabkan pengambilan keputusan yang buruk, kegagalan tugas, dan potensi bahaya keselamatan.

Jangkauan Cakupan Terbatas dan Konsekuensinya. Keterbatasan signifikan lainnya dari sistem sensorik agen yang berwujud adalah cakupannya. Agen hanya dapat merasakan apa yang berada dalam jangkauan deteksi sensornya. Misalnya, kamera mungkin memiliki bidang pandang yang terbatas, atau mikrofon mungkin hanya menangkap suara dalam radius tertentu. Di banyak lingkungan yang besar atau kompleks, cakupan terbatas tersebut dapat sangat membatasi kemampuan agen untuk mengumpulkan data komprehensif tentang lingkungannya.

Jangkauan terbatas berarti bahwa agen secara efektif "buta" terhadap apa pun di luar persepsi langsungnya. Dalam lingkungan dinamis, hal ini dapat menyebabkan kesenjangan informasi yang menghambat kesadaran situasional agen. Misalnya, dalam lingkungan yang bergerak cepat seperti lalu lintas, kendaraan otonom dengan jangkauan sensor terbatas mungkin gagal mendeteksi mobil atau pejalan kaki yang mendekat dari titik buta hingga terlambat untuk merespons. Seperti yang ditunjukkan pada Gambar 3.1, seseorang sedang menyeberang jalan tetapi terhalang pandangannya oleh kendaraan merah. Jika kendaraan biru

terus bergerak maju, ada risiko tabrakan yang tinggi. Hal ini dapat menyebabkan kecelakaan, membuat sistem tidak dapat diandalkan dan tidak aman dalam aplikasi dunia nyata.



Gambar 3.1 Ilustrasi rentang cakupan terbatas

Keterbatasan Adaptabilitas Lingkungan

Agen yang diwujudkan biasanya bergantung pada interaksi khusus dengan lingkungan fisiknya, yang dapat secara signifikan membatasi adaptabilitasnya ketika menghadapi perubahan dinamis dan lingkungan yang kompleks. Pertama, ketika menanggapi lingkungan yang berkembang pesat, kecepatan reaksi dan kemampuan adaptasi agen yang berwujud sering kali terbatas. Kemampuan mereka untuk menyesuaikan diri dengan perubahan lingkungan relatif lambat, terutama dalam skenario yang kompleks dan cepat berubah di mana satu agen yang berwujud mungkin kesulitan untuk memenuhi tuntutan tugas. Misalnya, dalam aplikasi kendaraan otonom, agen yang berwujud harus terus-menerus merasakan perubahan lingkungan dan membuat keputusan secara waktu nyata. Jika kecepatan perubahan ini melebihi kemampuan sensorik dan pengambilan keputusan agen, hal itu dapat menyebabkan kesalahan penilaian atau tindakan yang tertunda. Hal ini terutama terlihat dalam konteks yang sangat dinamis, seperti di lokasi kecelakaan lalu lintas, di mana perubahan yang cepat dan tidak terduga terjadi. Dalam kasus seperti itu, kemampuan adaptasi satu agen yang berwujud mungkin tidak mencukupi, sehingga membatasi kemampuannya untuk dengan cepat memodifikasi strateginya sebagai respons terhadap perubahan mendadak di lingkungan.

Kedua, agen yang berwujud sering dirancang dengan mempertimbangkan tugas dan lingkungan tertentu, yang membatasi fleksibilitas operasional mereka dalam konteks yang lebih luas. Sebagai contoh, robot yang dirancang untuk lingkungan industri mungkin dioptimalkan untuk lantai pabrik yang datar dan mudah diprediksi, di mana ia bekerja secara efisien. Namun, ketika digunakan di medan yang lebih tidak teratur atau menantang, seperti hutan atau daerah pegunungan, kinerjanya dapat menurun secara signifikan. Ketergantungan pada kondisi lingkungan tertentu ini mempersempit ruang lingkup aplikasi agen, membuatnya kurang serbaguna di berbagai lingkungan. Lebih lanjut, kurangnya generalisasi pada agen yang berwujud sangat menonjol dalam pengaturan multi-tugas.

Banyak agen yang berwujud dirancang untuk unggul dalam tugas-tugas tertentu, tetapi ketika sifat tugas berubah atau mencakup berbagai domain, keterbatasannya menjadi lebih

jelas. Misalnya, robot dalam manufaktur industri mungkin sangat mahir dalam merakit komponen tertentu. Namun, ketika ditugaskan untuk menangani material baru atau melakukan operasi yang tidak dikenal, mereka sering kali memerlukan pemrograman ulang atau modifikasi perangkat keras untuk mengakomodasi tuntutan baru. Kurangnya kemampuan beradaptasi terhadap tugas-tugas baru ini secara signifikan menghambat efektivitas mereka dalam menangani tantangan lintas domain yang kompleks.

Kapasitas Pengambilan Keputusan dan Penalaran yang Terbatas

Dibandingkan dengan SEAI, agen tunggal yang terwujud memiliki kemampuan pengambilan keputusan dan penalaran yang terbatas. Agen tersebut hanya dapat mengandalkan sensor dan daya komputasinya sendiri untuk merasakan lingkungan dan membuat penilaian, tanpa kemampuan untuk berbagi informasi atau bekerja sama dengan agen lain. Hal ini sering kali menyebabkan informasi yang tidak lengkap atau keputusan yang salah.

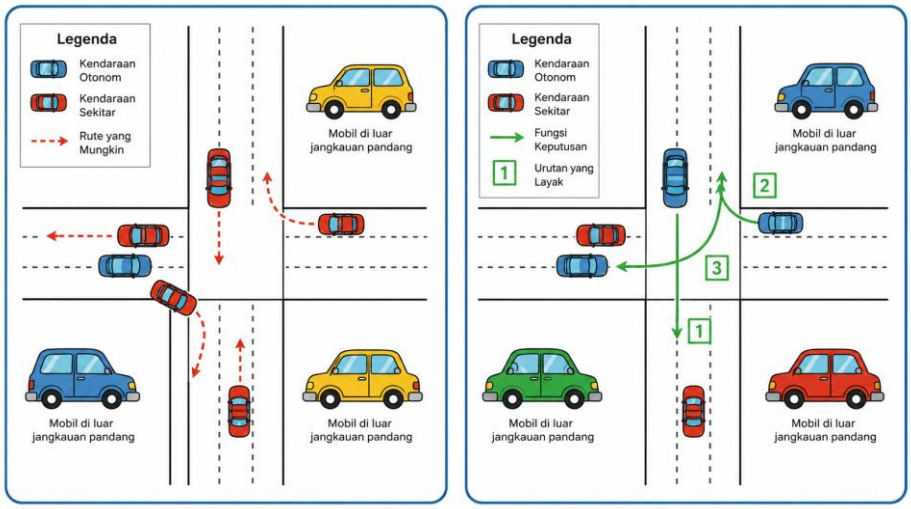
Kemampuan Pengambilan Keputusan yang Terbatas

Agen dapat berinteraksi dengan lingkungan eksternal melalui sensor dan membuat keputusan secara waktu nyata berdasarkan apa yang mereka rasakan. Namun, kemampuan pengambilan keputusan agen berwujud sering kali bergantung pada persepsi dan umpan balik lingkungan. Ketika mereka kekurangan pandangan global, kinerja mereka di lingkungan yang dinamis dan kompleks menjadi terbatas. Tanpa pandangan global, agen berwujud dapat jatuh ke dalam perangkap "optimum lokal" selama pengambilan keputusan. Hanya mengandalkan persepsi lokal dapat menyebabkan agen melewatkan jalur atau strategi terbaik secara keseluruhan. Misalnya, dalam tugas eksplorasi robot, satu robot dapat memilih jalur atau sumber daya terdekat, tetapi mengabaikan cara terbaik untuk menyelesaikan tugas secara keseluruhan. Batasan semacam ini yang disebabkan oleh penginderaan lokal sering kali membuat agen berkinerja lebih buruk daripada sistem dengan informasi global.

Selain itu, pengambilan keputusan kelompok dapat menghasilkan peningkatan pengetahuan yang lebih baik. Ketika beberapa agen berwujud bekerja bersama, mereka dapat belajar dari pengalaman satu sama lain untuk membuat keputusan yang lebih baik dan meningkatkan kemampuan sistem untuk merespons. Misalnya, jika satu agen mengalami kesulitan di lingkungan tertentu, ia dapat memperoleh bantuan atau strategi dari agen lain dan menghindari pengulangan kesalahan. Berbagi pengalaman membantu kelompok meningkatkan perilakunya selangkah demi selangkah dan beradaptasi dengan lingkungan dan tugas yang berbeda. Seiring agen berinteraksi lebih banyak, masing-masing dapat belajar dari yang lain dan meningkatkan diri, beralih dari "kecerdasan individu" ke "kecerdasan kolektif."

Selain itu, banyak tugas melibatkan banyak tujuan, variabel, dan lingkungan yang berubah. Saat membuat keputusan, agen individu mungkin salah menilai situasi karena mereka tidak dapat merasakan risiko atau peluang yang jauh. Misalnya, seperti yang ditunjukkan pada Gambar 3.2, ketika mobil otonom mendekati persimpangan, ia harus membuat keputusan terbaik untuk melewati dengan aman dan juga memberi jalan kepada mobil lain di dekatnya. Dalam kasus ini, rencana setiap mobil otonom mungkin saling bertentangan. Misalnya, mobil hijau ingin berbelok ke kiri. Jika tidak ada koordinasi yang tepat,

mobil tersebut mungkin menabrak mobil lain dan menghentikan semua mobil untuk melewati persimpangan dengan aman dan lancar.



Gambar 3.2 Contoh keterbatasan kemampuan pengambilan keputusan

Keterbatasan Kemampuan Penalaran pada Tugas-Tugas Kompleks

Dalam banyak situasi kehidupan nyata, agen yang berwujud tidak hanya dituntut untuk menyelesaikan satu tugas tunggal, tetapi juga harus beroperasi dalam lingkungan dinamis yang melibatkan koordinasi dan integrasi berbagai sub-tugas. Misalnya, tugas-tugas seperti penyelamatan bencana, mengemudi sendiri, atau pengiriman cerdas mencakup hal-hal seperti merasakan lingkungan, merencanakan jalur, bereaksi secara waktu nyata, dan menyeimbangkan berbagai tujuan. Semua ini merupakan bagian dari pengaturan multi-tugas yang kompleks. Namun, satu agen yang berwujud sering kali kekurangan kemampuan penalaran yang dibutuhkan untuk tugas-tugas kompleks. Penalaran bukan hanya tentang merasakan. Ini juga melibatkan pemahaman tugas, memprediksi apa yang mungkin terjadi, dan menyesuaikan rencana ketika keadaan berubah.

Ketika terlalu banyak informasi atau terlalu banyak aturan, agen dapat menjadi bingung. Agen mungkin tidak melihat gambaran lengkap atau membuat keputusan terbaik. Misalnya, ketika memilih di antara berbagai tujuan atau menangani konflik, agen mungkin memilih opsi terdekat atau termudah daripada apa yang terbaik untuk tugas secara keseluruhan. Keterbatasan kemampuan penalaran ini menyulitkan agen untuk berkinerja baik dalam lingkungan yang kompleks dan dinamis. Selain itu, agen berwujud (embodied agents) mengalami kesulitan dengan rantai tugas yang panjang dan kompleks. Banyak tugas dalam kehidupan nyata memiliki langkah-langkah yang saling bergantung.

Setiap langkah harus dilakukan dengan benar agar langkah selanjutnya dapat dimulai. Misalnya, dalam pekerjaan pengiriman, langkah-langkahnya meliputi pengemasan, perencanaan rute, dan pengantaran barang ke tempat yang tepat. Langkah-langkah ini saling terkait. Satu agen sering kali melakukan tugas dalam urutan sederhana dan tidak dapat menangani masalah yang memutus rantai tersebut. Jika satu langkah salah, agen tersebut

mungkin tidak tahu cara memperbaikinya, dan seluruh pekerjaan mungkin gagal atau melambat. Singkatnya, ketika melakukan tugas yang sulit dan multi-langkah, agen berwujud dibatasi oleh kemampuan penalaran mereka yang lemah. Hal ini menyulitkan mereka untuk membuat rencana yang baik atau bekerja dengan baik dalam situasi baru atau sulit. Penalaran yang lebih baik diperlukan jika kita ingin agen-agen ini menjadi lebih cerdas dan lebih bermanfaat dalam kehidupan nyata.

Ketahanan Operasional yang Terbatas

Dalam lingkungan dunia nyata yang kompleks dan dinamis, agen berwujud harus memiliki ketahanan operasional yang kuat untuk memastikan bahwa mereka dapat berinteraksi secara terus menerus dan andal dengan dunia fisik. Ketahanan operasional tidak hanya berarti bahwa sistem dapat terus berjalan, tetapi juga bahwa sistem tersebut dapat dengan cepat mendeteksi, membuat keputusan yang tepat, dan melakukan tugas dengan baik dalam kondisi yang tidak pasti dan bertekanan tinggi. Pertama, mencapai operasi yang tangguh membutuhkan pendeteksian yang cepat dan pemahaman yang tepat tentang lingkungan. Agen yang diwujudkan perlu memproses sejumlah besar data berdimensi tinggi dari sensor multi-modal, seperti penglihatan, pendengaran, dan sentuhan, dalam waktu yang sangat singkat.

Namun, satu agen memiliki daya komputasi dan kecepatan pemrosesan yang terbatas. Dalam lingkungan yang dinamis, sering kali sulit bagi satu agen untuk memberikan respons yang stabil dan efisien. Hal ini dapat menyebabkan penundaan, kesalahan dalam persepsi, atau keputusan yang salah, yang mengurangi kendali agen atas lingkungan. Kedua, satu agen berwujud juga menghadapi tantangan serius dalam penyimpanan dan komunikasi data. Tugas-tugas kompleks biasanya melibatkan aliran data berkelanjutan, seperti rangkaian gambar, suara, atau umpan balik fisik. Tetapi penyimpanan dan bandwidth dari satu agen biasanya sangat terbatas. Batasan perangkat keras ini dapat menyebabkan data penting hilang atau tertunda selama pengumpulan, transmisi, atau pemrosesan. Hal ini akan memengaruhi pengambilan keputusan dan mengurangi keandalan seluruh sistem.

Selain keterbatasan penginderaan dan komunikasi, ketahanan operasional juga dipengaruhi oleh kebisingan eksternal, pergantian tugas yang sering, dan penggunaan energi. Dalam kasus penggunaan nyata, ketidakpastian dan perubahan lingkungan membutuhkan kemampuan adaptasi yang lebih tinggi. Hal ini terutama berlaku untuk tugas-tugas seperti penyelamatan bencana, kendaraan otonom, atau operasi industri, di mana pekerjaan berkelanjutan dan toleransi kesalahan yang tinggi diperlukan. Dalam tugas-tugas ini, keterbatasan ketahanan agen tunggal lebih jelas terlihat. Oleh karena itu, untuk meningkatkan ketahanan operasional dalam kondisi dunia nyata, kita tidak hanya perlu mengoptimalkan persepsi dan pengambilan keputusan, tetapi juga merancang solusi tingkat sistem yang lebih baik. Ini termasuk penjadwalan data yang lebih efisien, desain toleransi kesalahan, dan kerja sama antar beberapa agen. Perubahan ini dapat membantu seluruh sistem menjadi lebih stabil dan lebih baik dalam menangani berbagai lingkungan.

Secara keseluruhan, agen berwujud tunggal menghadapi banyak keterbatasan dalam lingkungan dunia nyata yang kompleks. Dalam hal persepsi dan adaptasi, ia bergantung pada

sensor multimodal untuk memahami lingkungan, tetapi sering kali mengalami akurasi rendah, cakupan terbatas, dan kurangnya jenis sensor. Hal ini memudahkan agen untuk melewatkan informasi penting atau merespons terlalu lambat, terutama dalam tugas-tugas yang membutuhkan kesadaran situasional yang tinggi dan reaksi cepat. Selain itu, karena sebagian besar agen berwujud dirancang untuk tugas atau lingkungan tertentu, mereka memiliki kemampuan adaptasi yang buruk ketika tugas atau lingkungan berubah. Hal ini menyulitkan mereka untuk bekerja dengan baik di berbagai domain atau menangani banyak tugas. Dalam pengambilan keputusan dan penalaran, agen berwujud tunggal tidak memiliki akses ke informasi global dan tidak dapat berkoordinasi dengan yang lain.

Akibatnya, ia mungkin jatuh ke dalam perangkap "optimum lokal" dan membuat pilihan yang buruk. Ia juga kekurangan kemampuan penalaran untuk menangani tugas-tugas yang mencakup banyak langkah yang saling terkait. Ketika terjadi kesalahan, agen mungkin tidak tahu bagaimana cara menyesuaikan diri. Pada tingkat operasional, keterbatasan daya komputasi, penyimpanan, dan bandwidth komunikasi menyulitkan pemrosesan data sensor dalam jumlah besar secara waktu nyata. Hal ini dapat mengurangi stabilitas dan akurasi tindakan agen. Oleh karena itu, untuk meningkatkan tingkat kecerdasan keseluruhan agen berwujud, perlu diperkenalkan AI yang diwujudkan secara berkelompok (swarm Embodied AI). Hal ini dapat membantu meningkatkan sistem melalui persepsi kolaboratif, penalaran kolaboratif, dan bentuk koordinasi lainnya.

3.3 TEKNOLOGI UTAMA AI BERWUJUD KAWANAN

AI Berwujud Kawanan menggabungkan kekuatan kecerdasan kawanan (swarm intelligence) dan AI Berbasis Wujud (Embodied AI). AI ini menunjukkan kemampuan yang kuat dalam persepsi, pengambilan keputusan, dan adaptasi lingkungan. Tidak seperti agen tunggal, sistem kawanan menggunakan interaksi antara banyak agen untuk mencapai persepsi terdistribusi, pengaturan diri, dan optimasi strategi global dalam tugas-tugas kompleks. Di alam, sistem kolektif seperti koloni semut dan kawanan lebah menunjukkan pengambilan keputusan kolektif yang kuat, pengaturan diri, dan adaptasi dinamis. Contoh-contoh alami ini memberikan ide-ide yang berguna untuk merancang sistem buatan. Dengan memahami mekanisme kunci di balik perilaku kelompok ini, kita dapat membangun sistem kawanan yang lebih cerdas, lebih efisien, dan lebih tangguh.

Bagian ini memperkenalkan teknologi utama AI Berwujud Kawanan. Bagian ini berfokus pada tiga mekanisme penting: pengambilan keputusan kolektif, pengaturan dan homeostasis kolektif, dan pola periodik kolektif. Bagian-bagian ini menjelaskan bagaimana sistem alami bekerja dan bagaimana sistem tersebut menginspirasi desain algoritma. Selain itu, bagian ini menyajikan Optimasi Kawanan Partikel (PSO), algoritma terkenal yang belajar dari perilaku kelompok alami. Algoritma ini menunjukkan bagaimana aturan lokal dan umpan balik dapat membantu menyelesaikan masalah kompleks. Dengan menggabungkan ide dan metode ini, AI *Swarm Embodied* menjadi cara ampuh untuk menangani tugas-tugas kompleks di lingkungan yang berubah.

Kecerdasan Kawanan

Kecerdasan swarm adalah bidang fundamental dalam studi sistem multi-agen terdistribusi, di mana perilaku global yang kompleks muncul dari interaksi lokal sederhana dan tindakan individu. Contoh klasik kecerdasan kawanan di alam meliputi pencarian makanan semut, kawanan burung, dan pembangunan sarang lebah. Sistem alami ini menunjukkan kemampuan adaptasi, ketahanan, dan skalabilitas yang luar biasa, menawarkan inspirasi berharga untuk merancang sistem teknis. Namun, model tradisional kecerdasan kawanan sering kali bergantung pada abstraksi matematika dan kerangka kerja teoretis, yang mungkin tidak sepenuhnya memperhitungkan tantangan yang dihadapi dalam implementasi dunia nyata. Faktor-faktor seperti sifat fisik agen, kendala interaksi spasial, dan kompleksitas lingkungan dapat secara signifikan memengaruhi kinerja praktis model teoretis ini.

Untuk mengatasi tantangan ini, konsep AI Berwujud Kawanan (Swarm Embodied Artificial Intelligence/SEAI) telah diperkenalkan. SEAI menekankan pentingnya perwujudan fisik agen dan interaksi mereka dengan lingkungan dan dengan agen lain, memungkinkan koordinasi dan pengambilan keputusan yang lebih efektif dalam lingkungan yang dinamis dan kompleks. Tidak seperti kecerdasan kawanan tradisional, SEAI lebih menekankan pada karakteristik fisik agen dan bagaimana sifat-sifat ini memengaruhi perilaku kolektif. Pendekatan ini sangat relevan dengan sistem robot di dunia nyata, di mana sifat perwujudan agen membuat perilaku kawanan lebih realistis dan dapat diterapkan. Dengan menetapkan prinsip-prinsip inti untuk penyelidikan ilmiah, SEAI memperluas bidang ini dengan meningkatkan kemampuan adaptasi tugas dan meningkatkan kepraktisan sistem yang diwujudkan, sehingga menawarkan kerangka kerja yang lebih komprehensif untuk memahami dan merancang perilaku kawanan di lingkungan fisik.

SEAI telah banyak dipelajari dalam tulisan masa lampau. Salah satu definisi yang paling tepat diberikan oleh Marc Millonas, yang karyanya menguraikan persyaratan spesifik dan prinsip operasional algoritma SEAI. Inti dari algoritma ini adalah penggunaan aturan interaksi lokal sederhana dan komputasi untuk menghasilkan perilaku kawanan yang kompleks dan terukur. Dalam sistem alami, SEAI sering menampilkan tiga ciri utama: adaptabilitas, ketahanan, dan skalabilitas, yang merupakan sifat ideal yang ingin ditiru SEAI dalam sistem buatan.

- *Adaptabilitas*: Kawanan menunjukkan kemampuan untuk menyesuaikan diri dengan lingkungan dinamis dan menangani berbagai tugas secara efektif.
- *Ketahanan:* Kawanan mempertahankan fungsi yang stabil bahkan dalam menghadapi gangguan, seperti kegagalan atau kehilangan agen individu.
- *Skalabilitas:* Kinerja kawanan tetap konsisten seiring perubahan jumlah agen atau kompleksitas tugas. Penambahan atau penghapusan agen tidak secara signifikan menurunkan kinerja sistem, asalkan jumlah agen tetap di atas massa kritis.

Di luar ciri-ciri yang diakui secara luas ini, penskalaan kinerja super-linier merupakan fitur yang menentukan dari sistem SEAI. Ketika ukuran kawanan yang diwujudkan melebihi massa kritis, kinerja sistem kolektifnya harus menunjukkan penskalaan super-linier. Dengan kata lain, kawanan yang lebih besar tidak hanya mencapai produktivitas keseluruhan yang lebih tinggi

tetapi juga melampaui jumlah sederhana dari kontribusi agen individu. Dalam sistem SEAI yang sebenarnya, interaksi antar agen memperkuat kemampuan keseluruhan sistem, menghasilkan hasil yang lebih besar daripada jumlah bagian-bagiannya. Contoh efek super-linier tersebut telah diamati pada koloni lebah madu, kawanan robot, dan sistem komputasi multiprosesor. Sinergi kolaboratif antar agen meningkatkan efisiensi setiap agen individu, mengubah seluruh kawanan menjadi sistem yang sangat efektif.

Mencapai tingkat skalabilitas super-linier ini membutuhkan desain algoritma kontrol lokal yang cermat untuk agen individual. Algoritma ini harus memastikan bahwa tidak hanya efisiensi keseluruhan kawanan meningkat seiring bertambahnya ukurannya, tetapi juga bahwa kinerja agen individual meningkat secara positif seiring dengan ukuran kawanan. Fenomena ini, yang dikenal sebagai efek kawanan terwujud (embodied swarm effect), menghubungkan metrik kinerja sistem dengan kemampuan keseluruhan kawanan dan jumlah agen. Dibandingkan dengan Kecerdasan Individual Terwujud (Embodied Individual Intelligence), SEAI bergantung pada lebih dari sekadar kemampuan sensorik dan motorik agen individual. SEAI memanfaatkan kolaborasi dan interaksi dinamis antar agen untuk menghasilkan perilaku kelompok yang jauh melebihi jumlah kontribusi individu. Hal ini membuat SEAI sangat cocok untuk memecahkan masalah kompleks dan berskala besar. Sebaliknya, kecerdasan individual terwujud hanya bergantung pada kemampuan agen tunggal, yang secara inheren terbatas dalam menangani tugas-tugas kompleks dan berskala besar. Akibatnya, skalabilitas muncul sebagai tolok ukur penting untuk mengevaluasi kecerdasan sejati sistem SEAI.

Agar sistem SEAI unggul dalam skalabilitas, sistem tersebut harus memenuhi dua persyaratan mendasar. Pertama, kemampuannya harus muncul dari perilaku kolaboratif antar agen, bukan dari kemampuan individu yang terisolasi. Jika suatu sistem berkinerja sama baiknya dengan hanya satu agen seperti halnya dengan banyak agen, perilakunya tidak dapat diklasifikasikan sebagai kecerdasan kawanan (swarm-intelligent), karena tidak memiliki esensi kolaboratif SEAI. Kedua, seiring bertambahnya ukuran kawanan, kinerja keseluruhan sistem harus meningkat secara signifikan, dan kontribusi rata-rata setiap agen harus tetap stabil atau bahkan meningkat. Sistem SEAI yang benar-benar skalabel harus mendapat manfaat dari ukuran kawanan yang lebih besar, menghindari kehilangan efisiensi seiring bertambahnya jumlah agen. Penting untuk ditekankan bahwa karakteristik ini harus terwujud seiring bertambahnya ukuran kawanan secara bertahap, dari satu agen hingga sejumlah agen yang praktis, daripada diekstrapolasi ke ukuran tak terbatas.

Semua sistem SEAI beroperasi di bawah kendala fisik yang membatasi skalabilitasnya. Dengan mematuhi prinsip-prinsip ini, sistem SEAI dapat secara efektif memanfaatkan kolaborasi agen dan interaksi dinamis untuk mencapai adaptabilitas dan efisiensi yang luar biasa dalam memecahkan tantangan kompleks. Kecerdasan kawanan (swarm intelligence) mengambil inspirasi dari perilaku kolektif yang diamati di alam, di mana kelompok besar agen sederhana berinteraksi secara lokal untuk menghasilkan hasil yang kompleks dan cerdas. Tidak seperti sistem terpusat, kawanan alami ini menunjukkan koordinasi, ketahanan, dan kemampuan beradaptasi yang terdesentralisasi, sering kali mengandalkan aturan sederhana dan komunikasi lokal untuk mencapai tujuan global.

Dari koloni serangga hingga kelompok hewan dan komunitas mikroba, sistem ini menunjukkan bagaimana kerja sama dan umpan balik dapat memecahkan masalah secara efisien, bahkan dalam kondisi yang tidak pasti atau berubah. Untuk lebih memahami dan menerapkan prinsip-prinsip alami ini, kita mulai dengan meninjau beberapa contoh representatif kecerdasan kawanan di dunia alami. Contoh-contoh ini memberikan dasar intuitif dan membantu kita mengidentifikasi pola-pola kunci. Kemudian kita menganalisis tiga mekanisme inti yang terinspirasi dari biologi yang ditemukan di seluruh sistem ini, yaitu pengambilan keputusan kolektif, regulasi dan homeostasis kolektif, dan pola periodik kolektif, dan membahas bagaimana mekanisme tersebut memandu desain algoritma kawanan buatan.

Contoh Klasik Kecerdasan Kawanan

Kecerdasan kawanan muncul di banyak sistem alami, di mana individu-individu sederhana berinteraksi secara lokal dan menciptakan perilaku kelompok yang kompleks. Contoh-contoh berikut menunjukkan bagaimana semut, lebah, ikan, burung, dan mamalia menggunakan aturan sederhana untuk mencapai koordinasi kelompok yang efisien. **Koloni Semut**. Koloni semut menunjukkan kecerdasan kelompok yang luar biasa, yang sering disebut sebagai kecerdasan kawanan. Salah satu perilaku yang paling menarik adalah kemampuan mereka untuk menemukan jalur terpendek ke sumber makanan menggunakan jejak feromon. Semut menyimpan feromon di sepanjang jalur mereka, dan semut berikutnya lebih cenderung mengikuti jejak dengan konsentrasi feromon yang lebih tinggi.

Seiring waktu, ketika semut memperkuat jalur yang lebih pendek dengan lebih banyak feromon, koloni secara kolektif mengidentifikasi rute optimal. Fenomena ini, yang disebut stigmergy, memungkinkan semut untuk memecahkan masalah kompleks tanpa koordinasi pusat. Riset oleh Dorigo dkk. menyoroti bagaimana perilaku semut menginspirasi algoritma Optimasi Koloni Semut (ACO), yang digunakan untuk memecahkan masalah optimasi kombinatorial. Contoh lain adalah perilaku kooperatif mereka dalam membangun sarang yang rumit. Spesies semut seperti Atta (semut pemotong daun) membangun ruang bawah tanah yang rumit untuk membudidayakan kebun jamur, sumber makanan utama mereka. Tugas ini melibatkan penggalian yang tepat, pengelolaan limbah, dan pembagian kerja, yang menunjukkan organisasi sosial yang terdesentralisasi tetapi sangat efisien.

Koloni Lebah Madu. Lebah menunjukkan kerja sama tim yang kuat ketika mereka perlu menemukan sarang baru. Beberapa lebah, yang disebut pengintai, terbang keluar untuk mencari tempat yang baik. Ketika mereka kembali, mereka melakukan tarian goyang untuk menunjukkan lokasi dan kualitas tempat tersebut. Lebah lain mengamati, mengunjungi tempat-tempat tersebut, dan juga menari jika mereka setuju. Proses ini menyebarkan informasi dalam kelompok. Seiring waktu, sebagian besar lebah mendukung satu tempat, dan koloni berpindah. Tidak ada pemimpin yang memberi perintah. Keputusan dibuat berdasarkan kesepakatan kelompok melalui tarian dan kunjungan. Lebah juga mengontrol suhu sarang. Dalam cuaca panas, mereka mengipas sayap mereka atau membawa air untuk mendinginkan sarang. Pada saat dingin, mereka berkumpul rapat untuk tetap hangat. Lebah di luar berganti posisi untuk menjaga keselamatan semua orang.

Kawanan Ikan. Ikan sering berenang berkelompok untuk melindungi diri. Setiap ikan mengamati beberapa tetangganya dan meniru kecepatan dan arah mereka. Perilaku sederhana ini menciptakan gerakan yang halus dan sinkron di seluruh kelompok. Ketika predator menyerang, kelompok ikan mengubah bentuknya dengan cepat. Gerakan cepat ini membingungkan predator, sehingga sulit untuk menangkap satu ikan. Ini disebut "efek kebingungan." Beberapa ikan dalam kelompok mungkin mengetahui tentang sumber makanan atau ancaman. Gerakan mereka dapat memengaruhi yang lain melalui interaksi lokal. Bahkan tanpa pemimpin, kelompok ikan dapat mengubah arah dengan cepat dan beradaptasi dengan lingkungan.

Kawanan Burung. Burung seperti jalak terbang dalam kelompok besar yang disebut murmurasi. Setiap burung mengamati enam atau tujuh tetangganya dan menyesuaikan penerbangannya berdasarkan mereka. Bersama-sama, kelompok tersebut bergerak sebagai satu kesatuan, menghindari tabrakan sambil menciptakan bentuk yang mengalir di langit. Pola-pola ini membantu burung merespons dengan cepat terhadap predator atau perubahan arah. Kawanan tetap bersama tanpa kendali pusat. Aturan lokal sudah cukup untuk koordinasi skala besar. Dalam migrasi burung, burung yang lebih tua memimpin burung yang lebih muda di sepanjang rute yang panjang. Burung terbang dalam bentuk V untuk menghemat energi. Burung yang berada di depan menghadapi hambatan udara paling besar, jadi mereka bergiliran. Kerja sama tim ini membantu kelompok tersebut melakukan perjalanan jauh dengan energi yang lebih sedikit.

Kawanan Mamalia. Wildebeest bergerak dalam kawanan besar untuk mencari air dan rumput selama perubahan musim. Beberapa individu merasakan area yang lebih baik dan yang lain mengikuti. Informasi menyebar melalui isyarat sederhana seperti gerakan atau suara. Pergerakan kelompok membantu mereka menghindari predator. Tetap bersama membuat predator lebih sulit menyerang satu hewan. Saat menyeberangi sungai atau daerah berbahaya, mereka bergerak bergelombang untuk mengurangi risiko. Gajah juga hidup dalam kelompok sosial. Betina tertua, sang matriark, memimpin kelompok tersebut. Dia mengingat jalur dan sumber air dari masa lalu. Gajah menggunakan suara dan sentuhan untuk tetap terhubung, bahkan dalam jarak jauh. Mereka saling menjaga dan bekerja sama untuk melindungi anak-anak mereka.

Sistem alami dengan karakteristik kolektif memberikan sumber inspirasi yang kaya dan beragam untuk merancang algoritma kecerdasan kawanan. Sistem ini biasanya terdiri dari beberapa komponen yang berinteraksi, yang mungkin berupa organisme hidup atau entitas non-hidup. Untuk mengambil inspirasi dari sistem yang kompleks tersebut dan menerapkannya pada desain algoritma, sangat penting untuk menggali dua pertanyaan mendasar: “Mengapa sistem ini bekerja secara efektif?” dan “Bagaimana cara kerjanya?” Dengan menjawab pertanyaan-pertanyaan ini, kita dapat menyaring mekanisme inti yang mendorong fungsionalitas sistem dan mengabstraksikannya menjadi fitur-fitur utama yang sesuai untuk desain algoritma. Bagian-bagian berikut akan menganalisis sifat dan fungsionalitas kolektif yang paling signifikan dari kecerdasan kawanan, mengilustrasikan bagaimana hal tersebut memandu pengembangan algoritma inovatif.

Pengambilan Keputusan Kolektif

Pengambilan keputusan kolektif mengacu pada proses di mana suatu sistem, melalui interaksi dan berbagi informasi di antara individu-individunya, sampai pada pilihan optimal dari beberapa alternatif. Misalnya, semut menggunakan jejak feromon untuk menandai jalur, memungkinkan koloni untuk secara bertahap berkumpul pada rute terpendek. Demikian pula, lebah melakukan "tarian goyang" untuk berkomunikasi dan secara kolektif memutuskan area mencari makan atau tempat bersarang terbaik. Fenomena ini menyoroti bagaimana perilaku lokal dalam sistem kompleks dapat diperkuat melalui interaksi, yang pada akhirnya menghasilkan perilaku global, terkoordinasi, dan makroskopis. Dalam sistem pengambilan keputusan kolektif, lingkaran umpan balik positif memainkan peran sentral. Mekanisme ini beroperasi berdasarkan prinsip bahwa begitu perilaku lokal tertentu dimulai, perilaku tersebut diperkuat dan disebarkan melalui sistem, mendorong lebih banyak individu untuk mengadopsi perilaku yang sama. Misalnya, ketika semut mencari makan, individu secara acak menjelajahi lingkungannya.

Setelah menemukan makanan, semut meninggalkan jejak feromon di sepanjang jalurnya. Semut lain, yang mendeteksi feromon ini, lebih cenderung mengikuti jejak yang ditandai, sehingga semakin memperkuatnya. Seiring waktu, jalur dengan konsentrasi feromon yang lebih tinggi menarik lebih banyak semut, yang menyebabkan munculnya perilaku kolektif di mana koloni berkumpul pada jalur terpendek. Proses ini, di mana perilaku lokal diperkuat dan menyebar, tidak hanya mengoptimalkan tujuan keseluruhan sistem tetapi juga menggarisbawahi peran penting kontribusi individu terhadap koordinasi kelompok. Di luar mekanisme umpan balik, heterogenitas dan sifat terdistribusi individu dalam sistem pengambilan keputusan kolektif patut mendapat perhatian. Dalam sebagian besar sistem alami, heterogenitas individu secara signifikan meningkatkan stabilitas dan efisiensi pengambilan keputusan kolektif. Misalnya, dalam koloni lebah, individu tertentu, yang dikenal sebagai "pengintai," mengambil peran proaktif dalam menjelajahi lokasi sarang potensial selama relokasi. Sementara itu, lebah lain terutama mengevaluasi informasi yang diberikan oleh pengintai untuk membuat keputusan yang tepat. Pembagian kerja ini, yang muncul dari perbedaan individu, meningkatkan efisiensi eksplorasi sumber daya dan mengurangi risiko bias kelompok terhadap satu sumber informasi.

Sifat terdistribusi dari sistem pengambilan keputusan kolektif lebih lanjut memberikan sistem ini ketahanan dan kemampuan adaptasi yang luar biasa. Tidak seperti proses pengambilan keputusan terpusat, sistem terdistribusi memungkinkan individu untuk membuat keputusan independen berdasarkan isyarat lingkungan lokal dan kondisi mereka sendiri. Desentralisasi ini memastikan bahwa sistem dapat mempertahankan fungsionalitas keseluruhannya bahkan ketika beberapa individu gagal. Sifat terdistribusi tersebut memiliki implikasi signifikan untuk desain sistem kecerdasan buatan, khususnya dalam pengendalian kawanan robot skala besar dan optimasi jaringan terdesentralisasi.

Regulasi Kolektif dan Homeostasis

Salah satu ciri khas sistem kompleks, khususnya organisme biologis, adalah kemampuannya untuk mempertahankan variabel-variabel kritis tertentu dalam kisaran yang

layak dan efisien melalui regulasi kolektif. Fenomena ini, yang disebut homeostasis kolektif, memastikan bahwa sistem dapat berfungsi secara efektif dalam lingkungan yang dinamis, menghindari ketidakstabilan atau penipisan sumber daya. Di alam, hewan sosial sering menunjukkan contoh luar biasa dari kemampuan ini, di mana kerja sama antar individu mengarah pada pemanfaatan dan distribusi sumber daya yang efisien, mencapai homeostasis pada tingkat populasi. Inti dari sistem regulasi kolektif terletak pada lingkaran umpan balik negatif. Mekanisme ini beroperasi dengan melawan penyimpangan dari titik target yang telah ditetapkan, sehingga memulihkan keseimbangan. Ketika variabel sistem menjauh dari kisaran yang diinginkan, lingkaran umpan balik negatif memicu respons korektif untuk mengembalikan variabel tersebut ke posisi semula. Misalnya, pengaturan suhu dalam koloni lebah berfungsi sebagai contoh yang sangat penting. Lebah pekerja menggunakan perilaku seperti getaran sayap untuk mendinginkan sarang atau berkumpul bersama untuk menghemat panas, mempertahankan kisaran suhu yang optimal untuk perkembangan larva.

Ketika suhu sarang menjadi terlalu tinggi atau terlalu rendah, perilaku ini disesuaikan secara dinamis untuk memulihkan keseimbangan. Bentuk pengaturan ini berakar pada perilaku individu sederhana yang, ketika digabungkan, memungkinkan sistem secara keseluruhan untuk mempertahankan kendali lingkungan. Hasil yang muncul menunjukkan kekuatan interaksi lokal dalam menghasilkan stabilitas sistem yang terkoordinasi dan menyeluruh. Kekuatan utama dari sistem pengaturan kolektif adalah kemampuan adaptasi dan ketahanannya. Kemampuan adaptasi mengacu pada kemampuan sistem untuk menyesuaikan diri secara dinamis terhadap perubahan lingkungan, sementara ketahanan menunjukkan kapasitasnya untuk mempertahankan fungsionalitas meskipun terjadi gangguan internal atau eksternal. Sifat-sifat ini sangat jelas terlihat dalam sistem alami. Misalnya, ketika suatu sumber daya menjadi langka, semut atau lebah dengan cepat mengalihkan upaya mereka untuk menemukan dan memanfaatkan sumber alternatif, memastikan kelangsungan hidup koloni.

Demikian pula, jika individu tertentu dikeluarkan atau dinonaktifkan, anggota yang tersisa mengkompensasi kehilangan tersebut, sehingga mempertahankan fungsi keseluruhan sistem. Karakteristik ini tidak hanya menarik dari perspektif biologis tetapi juga sangat relevan dengan aplikasi teknologi dan teknik. Memahami bagaimana sistem alami mencapai kemampuan beradaptasi dan ketahanan dapat memberikan informasi untuk desain sistem buatan, seperti jaringan terdistribusi atau kawanan robot. Dalam bidang teknik dan kecerdasan buatan, salah satu tantangan utama adalah merancang sistem terdistribusi yang dapat mengelola sumber daya secara efisien dan menjaga stabilitas.

Dengan mempelajari mekanisme pengaturan kolektif di alam, para ahli dapat mengembangkan algoritma yang meniru proses ini. Misalnya, robot kawanan dapat memanfaatkan prinsip-prinsip homeostasis kolektif untuk mengoordinasikan pergerakan robot individu untuk tugas-tugas seperti pencarian dan penyelamatan atau pemantauan lingkungan. Demikian pula, sistem jaringan cerdas dapat mengadopsi strategi alokasi sumber daya berbasis umpan balik yang terinspirasi oleh sistem alami untuk mengoptimalkan distribusi energi dan meningkatkan ketahanan terhadap kegagalan.

Pola Periodik Kolektif

Sistem kolektif sering menunjukkan dinamika perubahan keadaan yang rumit dari waktu ke waktu, kadang-kadang menimbulkan pola spasial yang menarik seperti gelombang melingkar atau spiral. Pola spasial-temporal ini bukan sekadar produk sampingan dari perubahan keadaan dalam sistem tetapi memainkan peran fungsional yang penting. Misalnya, mereka dapat memfasilitasi komputasi kolektif, memungkinkan penilaian diri terhadap keadaan kelompok (misalnya, memperkirakan ukuran kelompok lokal), dan berfungsi sebagai media untuk transmisi sinyal. Munculnya dan penyebaran pola-pola ini menjelaskan prinsip-prinsip dasar komunikasi, regulasi umpan balik, dan kopling dinamis dalam sistem kompleks.

Pembentukan pola periodik dalam sistem kolektif terutama didorong oleh dua mekanisme utama: penundaan komunikasi dan perubahan keadaan internal individu yang kompleks. Penundaan komunikasi: Penundaan komunikasi muncul ketika transmisi informasi antar individu membutuhkan waktu, menciptakan efek jeda dalam sistem. Misalnya, dalam kelompok serangga tertentu, individu bertukar informasi melalui sinyal kimia atau isyarat perilaku. Kecepatan perambatan sinyal-sinyal ini menentukan seberapa cepat kelompok dapat merespons secara keseluruhan. Ketika penundaan ini terakumulasi dalam sistem, hal itu dapat menyebabkan perilaku osilasi yang teratur. Misalnya, umpan balik yang tertunda dalam pensinyalan dapat menghasilkan gelombang ritmis yang merambat melalui kelompok, memengaruhi perilaku kolektif. Perubahan Keadaan Internal yang Kompleks: Osilasi periodik juga dapat dihasilkan dari dinamika internal individu yang rumit. Banyak hewan menunjukkan pola perilaku yang dipengaruhi oleh ritme fisiologis atau keadaan energi internal mereka. Osilasi internal ini dapat merambat melalui interaksi antar individu, yang pada akhirnya bermanifestasi sebagai pola periodik yang dapat diamati pada tingkat kelompok. Pola-pola tersebut mengatur perilaku individu sekaligus memengaruhi fungsionalitas dan koordinasi keseluruhan sistem.

Pola periodik, khususnya yang diekspresikan sebagai gelombang spasial dinamis, menawarkan berbagai manfaat fungsional dalam sistem kolektif. Pola-pola ini dapat berfungsi sebagai alat komputasi terdesentralisasi, di mana frekuensi atau fase gelombang spasial-temporal mengkodekan informasi tentang distribusi sumber daya atau kondisi lingkungan. Hal ini memungkinkan sistem untuk memproses tugas-tugas kompleks secara kolektif tanpa bergantung pada kontrol terpusat. Misalnya, dalam sistem osilator yang terhubung, gelombang dapat mewakili ketersediaan sumber daya atau perubahan lingkungan, memungkinkan kelompok untuk merespons secara efisien sebagai satu kesatuan yang terkoordinasi. Pola-pola ini juga memainkan peran penting dalam penilaian diri terhadap keadaan kelompok. Banyak hewan sosial bergantung pada informasi tentang ukuran dan kepadatan kelompok untuk proses pengambilan keputusan. Dalam beberapa kasus, kecepatan atau amplitudo perambatan gelombang dapat memberikan informasi lokal kepada individu tentang penggunaan sumber daya atau kepadatan kelompok, memungkinkan umpan balik dinamis yang membantu mengoptimalkan perilaku kelompok, seperti menyempurnakan jangkauan pencarian makanan atau mengarahkan kembali pergerakan.

Lebih lanjut, pola periodik dapat bertindak sebagai media yang efektif untuk transmisi sinyal. Spesies tertentu, seperti katak dan serangga, menghasilkan gelombang ritmis melalui suara atau getaran untuk menyampaikan informasi, seperti panggilan kawin atau peringatan bahaya. Osilasi ini meningkatkan jangkauan dan keandalan komunikasi melalui sinkronisasi, memperkuat sinyal dan memastikan sinyal tersebut mencapai audiens yang lebih besar. Studi tentang pola periodik kolektif menawarkan wawasan mendalam tentang perilaku sistem alami sekaligus menginspirasi inovasi dalam sistem buatan. Dalam robotika, meniru pola-pola ini dapat memungkinkan koordinasi terdesentralisasi dalam sistem kawanan skala besar, seperti yang digunakan dalam operasi pencarian dan penyelamatan atau pemantauan lingkungan. Demikian pula, pemahaman tentang dinamika periodik dalam sistem biologis memiliki implikasi untuk riset medis, termasuk studi tentang aritmia jantung dan osilasi saraf.

Optimasi Kawanan Partikel

Studi tentang kecerdasan kolektif alami menawarkan wawasan mendalam tentang bagaimana sistem terdistribusi dapat mencapai koordinasi yang efisien, kemampuan adaptasi yang luar biasa, dan optimasi holistik tanpa bergantung pada kontrol terpusat. Dengan mengungkap prinsip-prinsip fundamental yang mendasari perilaku kelompok di alam, kita dapat mengubah wawasan ini menjadi strategi berharga untuk merancang sistem buatan yang mampu mengatasi masalah kompleks. Optimasi Kawanan Partikel (PSO), algoritma optimasi yang terinspirasi dari alam, meniru interaksi antar individu (partikel) dan kesadaran lingkungan mereka. Dengan menyesuaikan posisi partikel secara dinamis dalam ruang pencarian, PSO secara efisien menavigasi lanskap pemecahan masalah yang kompleks. Pada bagian ini, kita akan mengeksplorasi prinsip-prinsip inti dari algoritma PSO dan memeriksa bagaimana algoritma ini mengambil inspirasi dari perilaku kolektif yang diamati di alam.

PSO pertama kali diperkenalkan oleh Kennedy dan Eberhart. Ini adalah teknik optimasi meta-heuristik yang terinspirasi oleh perilaku sosial yang diamati di alam, seperti perilaku hewan, serangga, dan manusia. Sejak awal kemunculannya, PSO telah mendapatkan pengakuan luas di kalangan ahli dan praktisi karena ketangguhan dan efisiensinya dalam memecahkan berbagai masalah optimasi. Desain PSO didasarkan pada wawasan kunci tentang perilaku sosial dan kognisi manusia: “Orang belajar memahami dunia dengan berbicara dengan orang lain tentangnya”. Pengamatan ini menginspirasi Kennedy dan Eberhart untuk mengembangkan kerangka kerja komputasi di mana populasi solusi kandidat meningkat secara iteratif melalui berbagi informasi dan penyesuaian dinamis di antara partikel.

Dalam PSO, setiap partikel mewakili solusi potensial dalam ruang masalah. Partikel mengoptimalkan posisi mereka dengan mengintegrasikan umpan balik dari pengalaman mereka sendiri dan informasi yang dibagikan oleh tetangga mereka. Meskipun setiap partikel beroperasi berdasarkan aturan sederhana, interaksi kolektif antar partikel sering menghasilkan munculnya solusi optimal global (meskipun tidak dijamin). Partikel menyiarkan posisi mereka saat ini ke tetangga dan menyesuaikan pergerakan mereka menggunakan gangguan acak dan kombinasi kecepatan mereka, posisi terbaik pribadi, dan posisi terbaik yang ditemukan oleh partikel tetangga. Saat algoritma berulang, kawanan tersebut berkumpul menuju wilayah di ruang pencarian yang berisi solusi berkualitas tinggi. Proses ini meniru perilaku sekumpulan

burung yang secara kolaboratif mencari makanan: dengan berkoordinasi dan berkomunikasi, kawanan tersebut secara bertahap menemukan sumber makanan. Demikian pula, PSO memungkinkan partikel untuk berinteraksi dan secara kolektif menemukan solusi optimal dalam ruang pencarian.

Secara khusus, kecepatan dan posisi setiap partikel dalam PSO diperbarui secara iteratif berdasarkan keadaan saat ini, posisi terbaik pribadi, dan posisi terbaik yang ditemukan oleh tetangganya. Misalkan v_i menunjukkan kecepatan partikel ke-i, x_i posisi saat ini, p_i posisi terbaik pribadi, dan p_g posisi terbaik yang ditemukan di lingkungannya (atau seluruh kawanan). Pembaruan untuk v_i dan x_i dalam algoritma PSO asli didefinisikan oleh persamaan berikut:

$$\begin{aligned} \mathbf{v}_i^{\mathrm{NEW}} &\leftarrow \mathbf{v}_i + \varphi_1 \otimes (\mathbf{p}_i - \mathbf{x}_i) + \varphi_2 \otimes (\mathbf{p}_g - \mathbf{x}_i), \\ \mathbf{x}_i^{\mathrm{NEW}} &\leftarrow \mathbf{x}_i + \mathbf{v}_i^{\mathrm{NEW}} \end{aligned} \tag{3.1}$$

di mana $\varphi_1 = c_1 R_1$ dan $\varphi_2 = c_2 R_2$, R_1, dan R_2 adalah fungsi independen yang menghasilkan vektor nilai acak yang terdistribusi secara seragam dalam rentang. c_1 dan c_2 adalah koefisien percepatan, sedangkan *simbol* $\otimes$ *menunjukkan* perkalian vektor elemen demi elemen. Persamaan 3.1 mendefinisikan kecepatan v_i suatu partikel sebagai kombinasi dari tiga komponen: istilah "momentum," "kognitif," dan "sosial": Istilah "momentum" v_i mewakili inersia partikel, yang memungkinkannya untuk mempertahankan lintasan sebelumnya. Istilah "kognitif" $\varphi_1 \otimes (\mathbf{p}_i - \mathbf{x}_i)$ mencerminkan kecenderungan partikel untuk kembali ke posisi terbaik yang telah ditemukannya. Istilah "sosial" $\varphi_2 \otimes (\mathbf{p}_g - \mathbf{x}_i)$ menangkap kecenderungan partikel untuk bergerak menuju posisi terbaik yang diidentifikasi oleh kawanan secara keseluruhan. Dalam praktiknya, keacakan diperkenalkan menggunakan Pembangkit Bilangan Acak (RNG).

Namun, pilihan RNG dapat secara signifikan memengaruhi kinerja algoritma, dan yang mengejutkan, generator terbaik tidak selalu menghasilkan hasil yang optimal. Komponen "sosial" memanfaatkan topologi lingkungan untuk mengatur kecepatan penyebaran informasi antar partikel. Topologi umum meliputi konfigurasi cincin, bintang, dan von Neumann. Topologi lingkungan yang lebih kecil, seperti von Neumann, sering kali berkinerja lebih baik pada masalah multimodal yang kompleks karena aliran informasi yang terbatas, sementara lingkungan yang lebih besar cenderung unggul dalam masalah unimodal yang lebih sederhana. Ketika algoritma PSO membatasi p_g ke lingkungan lokal, algoritma ini disebut sebagai lbest PSO. Sebaliknya, jika $p_{\mathbf{g}}$ dipilih dari seluruh kawanan tanpa batasan, itu dikenal sebagai gbest PSO.

Algoritma 1 memberikan gambaran umum tentang implementasi PSO dasar. Dalam algoritma tersebut, f (.) mewakili fungsi kebugaran, dan $max(p_{neighbors})$ mengambil posisi terbaik di antara personal best dalam lingkungan partikel ke-i. Perlu dicatat bahwa Algoritma 1 biasanya mengadopsi model pembaruan asinkron klasik, yang secara luas dianggap sebagai yang paling efisien. Namun, pembaruan sinkron juga dapat digunakan, terutama dalam skenario komputasi paralel. Dalam kasus seperti itu, semua nilai pi pertama-tama disimpan,

diikuti dengan pembaruan ke p_g, dan kemudian pembaruan simultan ke kecepatan dan posisi semua partikel.

Algoritma 1 Algoritma PSO, dengan asumsi maksimisasi

Hasilkan populasi awal secara acak
ulangi
  untuk setiap partikel i **lakukan**
    jika $f(x_i) > f(p_i)$ **maka**
      $p_i \leftarrow x_i$;
      $p_g \leftarrow \max(p_{neighbors})$;
      Perbarui kecepatan v_i;
      Perbarui posisi p_i;
    akhir jika
  akhir untuk
sampai *kriteria penghentian terpenuhi*;

Banyak masalah optimasi dunia nyata pada dasarnya bersifat dinamis, membutuhkan algoritma yang dapat beradaptasi dengan perubahan solusi optimal dari waktu ke waktu. Misalnya, kondisi lalu lintas di kota terus berfluktuasi, apa yang dianggap sebagai rute terbaik pada suatu saat mungkin tidak lagi optimal tak lama kemudian. Tidak seperti optimasi statis, tujuan utama dalam lingkungan dinamis adalah untuk melacak dan menyesuaikan pergeseran optimal sedekat mungkin.

Particle Swarm Optimization (PSO) sangat cocok untuk optimasi dinamis karena konvergensi yang cepat dan fleksibilitas yang melekat. Melalui interaksi berkelanjutan antar partikel dalam swarm, PSO secara dinamis menyesuaikan posisi partikel, memungkinkan algoritma untuk merespons perubahan lingkungan secara efektif. Kemampuan adaptasi ini membuat PSO sangat menarik untuk memecahkan masalah dinamis. Namun, riset telah menunjukkan bahwa PSO standar harus diadaptasi untuk mengatasi tantangan tambahan yang ditimbulkan oleh lingkungan dinamis. Salah satu aplikasi PSO paling awal untuk optimasi dinamis dilakukan oleh Eberhart dan Shi. Mereka menggunakan PSO berbobot inersia untuk melacak optimum fungsi parabola unimodal yang nilai maksimumnya berubah secara berkala. Temuan mereka mengungkapkan bahwa, dalam kondisi tertentu, PSO dapat berkinerja setara atau bahkan melampaui algoritma evolusi. Mendeteksi perubahan dalam lingkungan dinamis dapat dicapai melalui beberapa metode. Salah satu pendekatan melibatkan penunjukan partikel penjaga selama setiap iterasi.

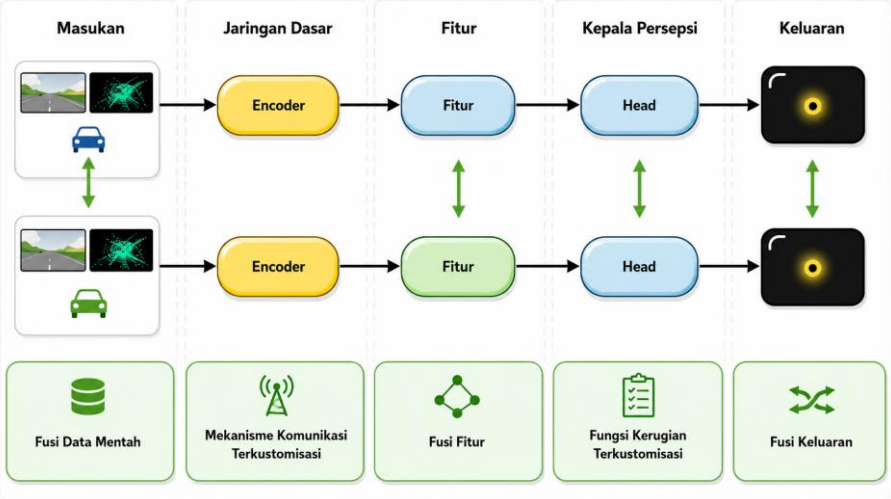
Dengan mengevaluasi kebugaran partikel penjaga dan membandingkannya dengan nilai sebelumnya, algoritma dapat mengidentifikasi perubahan lingkungan. Jika perubahan terdeteksi, kawanan akan diberi peringatan, memicu mekanisme respons yang sesuai. Metode sederhana lainnya adalah mengevaluasi kembali posisi terbaik global dan kebugaran partikel terbaik kedua untuk menentukan apakah telah terjadi perubahan. Untuk mengatasi masalah memori usang yang disebabkan oleh pergeseran lingkungan, berbagai strategi respons telah

diusulkan. Salah satu pendekatan umum melibatkan penggantian posisi terbaik historis setiap partikel secara berkala dengan posisi saat ini ketika perubahan terdeteksi. Hal ini memungkinkan partikel untuk membuang informasi usang dan hanya mengandalkan pengetahuan terbaru tentang lingkungan baru. Alternatifnya, pengacakan ulang sebagian dari kawanan telah disarankan untuk mempertahankan keragaman dan meningkatkan kemampuan kawanan untuk melacak optima baru.

PSO multi-kawanan, dikombinasikan dengan partikel bermuatan, juga telah menunjukkan potensi yang kuat. Dalam pendekatan ini, beberapa kawanan didistribusikan di berbagai puncak dalam ruang pencarian. Untuk mencegah konvergensi yang tumpang tindih, kawanan yang lebih lemah diacak ketika dua kawanan terlalu dekat. Selain itu, PSO multi-kawanan menggantikan partikel bermuatan dengan partikel kuantum, yang posisinya ditentukan secara probabilistik di sekitar penarik kawanan. Pendekatan ini sangat efektif dalam lingkungan dinamis multimodal yang kompleks di mana beberapa optima perlu dilacak secara bersamaan. Memperluas konsep ini, PSO berbasis spesies telah dikembangkan untuk mengidentifikasi dan mengikuti beberapa puncak dalam skenario dinamis. Dengan menggabungkan algoritma spesiasi, metode ini menggunakan "benih spesies" lokal untuk memandu partikel dalam radius yang ditentukan menuju optima lokal. Pendekatan ini mendorong kawanan untuk berkumpul pada beberapa solusi lokal daripada satu optimum global, sehingga sangat cocok untuk tugas optimasi multimodal di lingkungan dinamis.

3.4 KOMPONEN UTAMA AI BERWUJUD KAWANAN

AI berwujud kawanan mendorong persepsi yang lebih baik dan lebih holistik dengan memungkinkan banyak agen untuk berbagi informasi persepsi yang saling melengkapi. Tugas ini secara fundamental dapat mengatasi beberapa masalah yang sudah lama ada dalam persepsi agen tunggal, seperti oklusi. Metode dan sistem terkait memiliki potensi yang sangat besar dalam banyak aplikasi, termasuk multi-UAV (kendaraan udara tak berawak) untuk pencarian dan penyelamatan, otomatisasi dan pemetaan multi-robot, dan kolaborasi kendaraan-ke-kendaraan (V2V) dan kendaraan-ke-segala-sesuatu (V2X). Sistem persepsi kolaboratif umum terdiri dari data mentah, jaringan dasar, fitur, jaringan kolaborasi, dan keluaran. Berdasarkan tahap berbagi dan penggabungan data, tahapan persepsi kolaboratif secara luas dapat dipisahkan menjadi kolaborasi awal, menengah, dan akhir, seperti yang ditunjukkan pada Gambar 3.3:



Gambar 3.3 Tiga tahapan kolaborasi tipikal

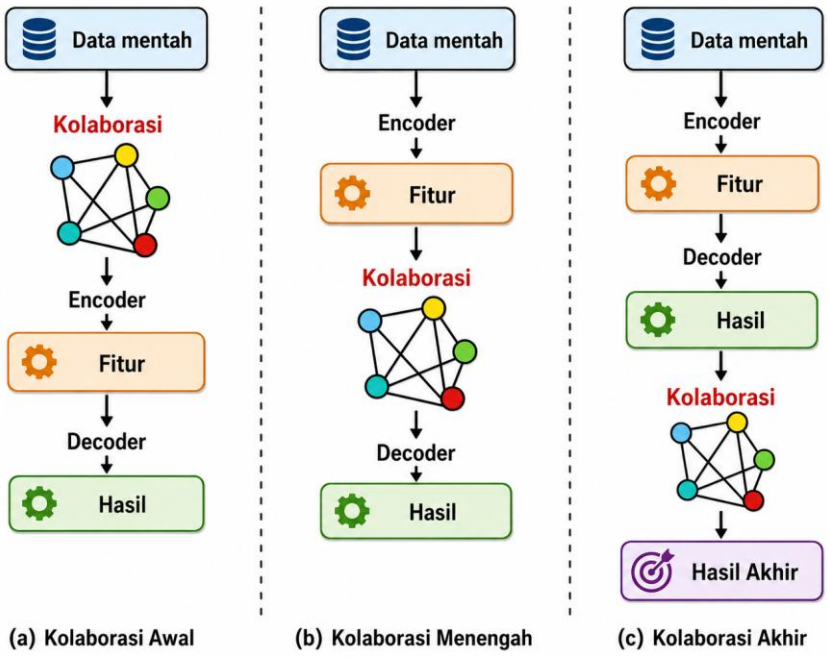
Kolaborasi Awal. Kolaborasi awal menggunakan fusi data mentah pada input jaringan kolaborasi, yang juga dikenal sebagai fusi tingkat data atau tingkat rendah. Agen ego menerima dan mengubah data sensor mentah dari agen lain, dan kemudian menggabungkan data yang telah diubah di dalam perangkat. Data mentah berisi informasi paling komprehensif dan deskripsi substansial dari agen. Akibatnya, kolaborasi awal secara fundamental dapat mengatasi masalah oklusi dan jangkauan jauh dalam persepsi individu dan meningkatkan kinerja hingga tingkat maksimal. Namun, kolaborasi awal bergantung pada bandwidth data yang tinggi, yang membuatnya sulit untuk mencapai komputasi tepi waktu nyata.

Kolaborasi Menengah. Mengingat bandwidth yang tinggi pada kolaborasi awal, beberapa karya mengusulkan metode persepsi kolaboratif menengah untuk menyeimbangkan trade-off kinerja-bandwidth. Dalam kolaborasi menengah, agen lain biasanya mentransfer fitur semantik mendalam ke agen ego. Agen ego menggabungkan fitur untuk membuat prediksi akhir. Kolaborasi menengah telah menjadi pilihan persepsi kolaboratif multi-agen yang paling populer untuk fleksibilitas. Namun, ekstraksi fitur sering menyebabkan kehilangan informasi dan redundansi informasi yang tidak perlu, yang memotivasi orang untuk mengeksplorasi strategi fusi fitur yang sesuai.

Kolaborasi Akhir. Kolaborasi tingkat akhir atau tingkat objek menggunakan fusi prediksi pada output jaringan. Setiap agen melatih jaringan secara individual dan berbagi output satu sama lain. Agen ego secara spasial mengubah output dan menggabungkan semua output setelah pasca-pemrosesan. Kolaborasi tingkat akhir lebih hemat bandwidth dan lebih sederhana daripada kolaborasi awal dan menengah. Namun, kolaborasi tingkat akhir juga memiliki keterbatasan. Karena output individual bisa bising dan tidak lengkap, kolaborasi tingkat akhir selalu memiliki kinerja persepsi terburuk.Banyak metode persepsi kolaboratif telah muncul baru-baru ini. Karena metode-metode ini didasarkan pada kerangka kerja dan tujuan yang berbeda, sangat penting untuk membangun taksonomi sistematis untuk membantu pembaca memahami bidang ini. Untuk tujuan ini, bagian ini meninjau pendekatan persepsi kolaboratif terbaru secara sistematis. Secara khusus, kami meninjau persepsi kolaboratif dari aspek metode untuk skenario ideal dan metode untuk skenario dunia nyata.

Yang pertama memperhatikan efisiensi dan kinerja kolaborasi, sedangkan yang terakhir lebih berfokus pada ketahanan dan keamanan kolaborasi.

Dalam sistem persepsi kolaboratif, para ahli biasanya merancang modul kolaborasi yang sesuai pada berbagai tahapan jaringan untuk meningkatkan efisiensi kolaborasi dan kinerja persepsi. Secara umum, tiga skema kolaborasi membutuhkan modul kolaborasi dasar untuk menggabungkan pengamatan multi-agen, seperti fusi data mentah pada tahap input (kolaborasi awal), fusi fitur pada tahap pemrosesan fitur (kolaborasi menengah), dan fusi output pada tahap output (kolaborasi akhir). Lebih lanjut, beberapa riset membangun mekanisme komunikasi pada tahap jaringan dasar untuk mengurangi bandwidth transmisi, dan memperkenalkan fungsi kerugian yang disesuaikan pada tahap kepala persepsi untuk memandu jaringan agar menangkap informasi yang bermanfaat. Pada sub-bagian ini, kami merangkum dan memperkenalkan modul kolaborasi dasar dari setiap tahap secara terpisah, seperti yang ditunjukkan pada Gambar 3.4.



Gambar 3.4 Ilustrasi modul kolaborasi umum

Kolaborasi Awal

Kolaborasi awal mengadopsi fusi data mentah pada tahap input. Karena awan titik tidak beraturan dan dapat diagregasi secara langsung, kolaborasi awal biasanya mengadopsi strategi fusi awan titik. Agen perlu merekonstruksi data yang diterima karena diambil pada posisi dan sudut yang berbeda. Dengan bertukar data LiDAR, lingkungan lokal dapat direkonstruksi secara intuitif dengan menggabungkan awan titik ke posisi fisiknya. Untuk merekonstruksi lingkungan lokal dengan memetakan awan titik ke posisi fisik, informasi tambahan dienkapsulasi ke dalam paket pertukaran. Paket tersebut harus terdiri dari informasi instalasi sensor LiDAR dan pembacaan GPS-nya, yang menentukan posisi titik pusat setiap frame awan titik. Pembacaan

IMU (unit pengukuran inersia) agen juga diperlukan karena merekam informasi offset agen selama bergerak.

Formulasi Masalah. Pertimbangkan N agen dalam suatu adegan. Setiap agen merepresentasikan rotasi dengan sudut yaw, pitch, dan roll masing-masing sebesar α, β, dan γ. Matriks rotasi R dihasilkan sebagai berikut:

$$R = R_z(\alpha)R_y(\beta)R_x(\gamma), \tag{3.2}$$

$$R_z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.3}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \tag{3.4}$$

$$R_x(\gamma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix} \tag{3.5}$$

di mana $R_z(\alpha), R_y(\beta)$, dan $R_x(\gamma)$ adalah tiga matriks rotasi dasar yang memutar vektor terhadap sumbu z-, y-, dan x= dalam ruang tiga dimensi.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R \\ Y_R \\ Z_R \end{bmatrix} \cup \begin{bmatrix} X'_T \\ Y'_T \\ Z'_T \end{bmatrix} \tag{3.6}$$

$$\begin{bmatrix} X'_T \\ Y'_T \\ Z'_T \end{bmatrix} = R \times \begin{bmatrix} X_T \\ Y_T \\ Z_T \end{bmatrix} + \begin{bmatrix} \Delta d_{x_T} \\ \Delta d_{y_T} \\ \Delta d_{z_T} \end{bmatrix} \tag{3.7}$$

Catatan: X, Y, dan Z dalam koordinat $[XYZ]$' merepresentasikan nilai spasial 3D dari setiap titik pada data awan titik LiDAR, sedangkan $[X'_T, Y'_T, Z'_T]'$ adalah awan titik pemancar setelah transformasi R diterapkan ke koordinat yang ditranslasikan dari agen pengirim. Transformasi dihitung dengan Persamaan 3.2, menggunakan perbedaan nilai IMU antara pemancar dan penerima. Ketika agen yang terhubung bertukar pesan, persepsi kooperatif menghasilkan bingkai baru dengan menggabungkan data sensor pemancar dan penerima menggunakan Persamaan 3.6, di mana kita memiliki himpunan semua koordinat yang sama dengan koordinat penerima yang digabungkan dengan koordinat dari pemancar. Namun, karena agen pengirim berada dalam keadaan yang berbeda dari penerima, kita harus menerapkan transformasi ke koordinat asli agar sesuai dengan keadaan agen penerima. Untuk mendapatkan keadaan yang benar untuk orientasi pemancar, kita menggunakan Persamaan 3.2.

Metode Unggulan. Sistem persepsi kolaboratif awal pertama Cooper memilih data LiDAR sebagai target fusi. Awan titik dapat dikompresi menjadi ukuran yang lebih kecil hanya dengan mengekstrak koordinat posisi dan nilai refleksi. Setelah interaksi antar agen, Cooper merekonstruksi awan titik yang diterima dengan matriks transformasi, dan kemudian menggabungkan himpunan awan titik ego untuk membuat prediksi akhir. Terinspirasi oleh

Cooper, Coop3D juga mengeksplorasi kolaborasi awal dan memperkenalkan metode fusi awan titik baru. Secara khusus, alih-alih menggunakan penggabungan, sistem Coop3D memanfaatkan transformasi spasial untuk menggabungkan data sensor. Selain itu, tidak seperti Cooper yang berbagi informasi antar agen di dalam pesawat, Coop3D mengusulkan sistem pusat untuk menggabungkan beberapa data sensor, yang memungkinkan amortisasi biaya sensor dan pemrosesan dalam kolaborasi.

Mekanisme Komunikasi

Fusi data mentah dalam kolaborasi awal memperluas cakupan agen ego, yang juga menyebabkan tekanan bandwidth yang tinggi. Untuk mengurangi masalah di atas, semakin banyak karya mengembangkan kolaborasi menengah. Metode kolaborasi perantara awal mengikuti mekanisme komunikasi serakah untuk memperoleh informasi sebanyak mungkin. Umumnya, mereka berbagi informasi dengan semua agen dalam jangkauan komunikasi dan memasukkan peta fitur lengkap yang dikompresi ke dalam pesan persepsi kolektif (CPM). Namun, karena kelangkaan fitur dan redundansi agen, komunikasi serakah dapat membuang bandwidth secara besar-besaran. Untuk mengatasi kesenjangan ini, beberapa riset menetapkan mekanisme komunikasi dinamis untuk memilih agen dan fitur.

Perumusan Masalah. Pertimbangkan N agen dalam suatu adegan. Misalkan $\mathcal{X}_i$ dan $\mathcal{Y}_i$ masing-masing adalah observasi dan supervisi persepsi dari agen ke-i. Tujuan persepsi kolaboratif adalah memaksimalkan kinerja persepsi semua agen sebagai fungsi dari total anggaran komunikasi B dan putaran komunikasi K; yaitu,

$$\begin{aligned} \xi_\Phi(B,K) \quad &= \arg\max_{\theta,\mathcal{P}} \sum_{i=1}^{N} g\left(\Phi_\theta\left(\mathcal{X}_i, \{\mathcal{P}_{i\to j}^{(K)}\}_{j=1}^{N}\right), \mathcal{Y}_i\right), \\ \text{s.t.} \quad & \sum_{k=1}^{K}\sum_{i=1}^{N}\left|\mathcal{P}_{i\to j}^{(k)}\right| \le B \end{aligned} \tag{3.8}$$

di mana g(·,·) adalah metrik evaluasi persepsi, Φ adalah jaringan persepsi dengan parameter latih θ, dan $P_{i\to j}^{(k)}$ adalah pesan yang dikirim dari agen ke-i ke agen ke-j pada putaran komunikasi ke-k. Perhatikan bahwa (i) ketika $B = K = 0$, tidak ada kolaborasi dan ξ_Φ (0, 0) mencerminkan kinerja persepsi agen tunggal; (ii) melalui pengoptimalan strategi komunikasi dan parameter jaringan, persepsi kolaboratif harus berkinerja baik secara konsisten pada bandwidth atau putaran komunikasi apa pun; dan (iii) Untuk komunikasi multi-putaran, setiap agen berfungsi sebagai pendukung (menawarkan pesan untuk membantu orang lain) dan pemohon (meminta pesan dari orang lain).

Metode Unggulan. Who2com memperkenalkan mekanisme komunikasi pertama yang dirancang untuk beroperasi di bawah kendala bandwidth, menggunakan proses jabat tangan tiga tahap. Ia menggunakan fungsi perhatian umum untuk menghitung skor kecocokan antar agen, memungkinkan pemilihan agen yang paling relevan, sehingga mengurangi penggunaan bandwidth. Berdasarkan hal ini, When2com memanfaatkan perhatian umum berskala untuk menentukan waktu optimal untuk komunikasi, memungkinkan agen ego untuk berinteraksi

dengan agen lain hanya ketika informasi tambahan diperlukan, sehingga menghemat sumber daya kolaborasi.

Selain memilih mitra komunikasi yang tepat, isi komunikasi juga sangat penting dalam meminimalkan tekanan bandwidth. FPV-RCNN adalah salah satu yang pertama kali mengusulkan strategi pemilihan fitur, menggunakan kepala deteksi untuk menghasilkan proposal dan memilih titik fitur dalam proposal ini. Modul pemilihan titik kunci ini mengurangi redundansi dalam fitur mendalam yang dibagikan dan meningkatkan kualitas proposal awal. Where2comm memperkenalkan mekanisme komunikasi sadar kepercayaan spasial baru yang lebih mengoptimalkan efisiensi komunikasi. Ia menggunakan peta kepercayaan spasial untuk menentukan fitur mana yang akan dibagikan dan dengan siapa akan berkomunikasi. Selama pemilihan fitur, Where2comm memilih dan mengirimkan hanya elemen spasial yang memenuhi kriteria kepercayaan tinggi dan diminta oleh agen lain. Dalam pemilihan agen, agen ego hanya berkomunikasi dengan mereka yang dapat menyediakan fitur yang diperlukan. Dengan berfokus pada area-area yang penting secara perseptual, Where2comm secara signifikan mengurangi penggunaan bandwidth dan meningkatkan efisiensi kolaborasi.

Kolaborasi Menengah

Modul fusi fitur memainkan peran penting dalam kolaborasi menengah. Setelah agen ego menerima CPM dari agen lain, ia dapat menggunakan berbagai strategi untuk menggabungkan fitur-fitur ini. Strategi fusi yang efektif dapat mengidentifikasi hubungan tersembunyi antar fitur dan meningkatkan kinerja jaringan persepsi. Berdasarkan pendekatan yang mendasarinya, metode fusi fitur dapat dikategorikan menjadi teknik tradisional, berbasis grafik, dan berbasis perhatian.

Fusi Fitur Tradisional

Pada fase awal riset persepsi kolaboratif, strategi tradisional seperti penggabungan, penjumlahan, dan pembobotan linier umumnya digunakan untuk menggabungkan fitur. Kolaborasi menengah biasanya menggunakan operasi invarian permutasi ini pada fitur-fitur mendalam, menghasilkan inferensi yang cepat karena kesederhanaannya. **Perumusan Masalah.** Pertimbangkan N agen dalam adegan. Misalkan *Oi* adalah peta fitur dari agen ke-i. Ketika agen penerima menerima peta fitur yang dibagikan dari agen pengirim, ia pertama-tama mengukur semua fitur baru yang terkandung dalam peta fitur tersebut, dengan membandingkannya dengan fitur-fitur dalam peta fiturnya sendiri. Berdasarkan pengukuran tersebut, penerima menerapkan bobot, X, pada peta fitur yang diterima sebelum fusi. Oleh karena itu, fusi tersebut menangkal efek negatif yang disebabkan oleh fitur-fitur lemah dalam peta fitur penerima.

$$\mathbf{F} = \left\{\mathbf{F}_3 \cup \max\{\mathbf{F}_1, \mathbf{F}_2 \times X\}\right\} \times Y \tag{3.9}$$

$$X = \begin{cases} S/(A_o/A), & S < a, \\ S/(A_o/A), & a \leq S < b, \\ K, & S \geq b \end{cases} \tag{3.10}$$

$$S = \|\mathbf{F}_1^i - \mathbf{F}_2^i\|/(W \times H) \tag{3.11}$$

di mana $\mathbf{F}_1$ dan $\mathbf{F}_2$ masing-masing adalah area tumpang tindih dari dua peta fitur yang akan digabungkan dari penerima dan pengirim; $\mathbf{F}_3$ adalah area tidak tumpang tindih dari peta fitur penerima; dan $\mathbf{F}$ adalah peta fitur gabungan yang dihasilkan. X adalah bobot yang diberikan pada peta fitur pengirim $\mathbf{F}_2$, sedangkan Y adalah parameter peningkatan fitur. di mana $S = \|\mathbf{F}_1^i - \mathbf{F}_2^i\|/(W \times H)$ dan $\|\mathbf{F}_1^i - \mathbf{F}_2^i\|$ adalah jarak L2 (Euclidean) antara dua peta fitur yang direpresentasikan oleh vektor $\mathbf{F}_1^i$ dan $\mathbf{F}_2^i$. W dan H adalah lebar dan tinggi area tumpang tindih dari dua peta fitur. A_o adalah ukuran area tumpang tindih, sedangkan A adalah ukuran seluruh peta fitur. Konstanta dalam persamaan di atas, misalnya a, b, k, diperoleh dari eksperimen intensif.

Metode Unggulan. Kerangka kerja persepsi kolaboratif menengah pertama F-Cooper mengekstrak fitur voxel tingkat rendah dan fitur spasial dalam. Berdasarkan fitur dua tingkat ini, F-Cooper mengusulkan dua strategi fusi fitur: Voxel Feature Fusion (VFF) dan Spatial Feature Fusion (SFF). Keduanya menggunakan maxout elemen demi elemen untuk menggabungkan fitur di wilayah yang tumpang tindih. Karena fitur voxel lebih dekat ke data mentah, VFF sama mampunya dengan metode fusi data mentah untuk deteksi objek dekat. Sementara itu, SFF juga memiliki keunggulannya. Terinspirasi oleh SENet, SFF memilih untuk memilih saluran parsial untuk mengurangi konsumsi waktu transmisi sambil mempertahankan presisi deteksi yang sebanding.

Mengingat F-Cooper mengabaikan pentingnya fitur dengan kepercayaan rendah, CoFF diusulkan untuk meningkatkan F-Cooper. CoFF memberi bobot pada fitur yang tumpang tindih dengan mengukur kesamaan dan area tumpang tindihnya. Semakin kecil kesamaan dan semakin besar jaraknya, semakin banyak informasi tambahan yang diberikan oleh fitur tetangga secara intuitif. Selain itu, parameter peningkatan ditambahkan untuk meningkatkan nilai fitur lemah. Eksperimen menunjukkan bahwa desain yang sederhana namun efisien membuat CoFF meningkatkan F-Cooper secara signifikan.

Meskipun sederhana, fusi tradisional belum ditinggalkan oleh metode-metode terbaru. Deteksi 3D berbasis kamera kolaboratif (CoCa3D) diusulkan untuk menunjukkan potensi kolaborasi dalam meningkatkan deteksi 3D berbasis kamera. Karena estimasi kedalaman merupakan hambatan utama deteksi 3D berbasis kamera, CoCa3D berisi estimasi kedalaman kolaboratif (Co-Depth) selain pembelajaran fitur kolaboratif (Co-FL). Dalam Co-Depth, agen tetangga hanya mengirimkan estimasi kedalaman dengan ketidakpastian rendah, dan agen ego memperbarui estimasi kedalaman dengan mempertimbangkan probabilitas kedalaman tampilan tunggal dan konsistensi multi-tampilan. Dalam Co-FL, agen mengirimkan elemen fitur dengan kepercayaan deteksi tinggi dan mengadopsi maksimum titik-per-titik non-parametrik sederhana untuk menggabungkan fitur. Hasil eksperimen menunjukkan bahwa CoCa3D membantu kamera mengungguli LiDAR dalam deteksi objek 3D.

Penggabungan Fitur Berbasis Graf

Meskipun penggabungan perantara tradisional sederhana, metode ini mengabaikan hubungan laten antar agen multi-agen dan gagal untuk menalar pesan dari pengirim ke penerima. Jaringan Neural Graf (GNN) memiliki kemampuan untuk menyebarkan dan menggabungkan pesan dari tetangga, dan karya-karya terbaru telah menunjukkan efektivitas

GNN pada persepsi dan mengemudi otonom. **Perumusan Masalah.** Misalkan X_i adalah awan titik 3D yang dikumpulkan oleh agen ke-i (total M agen), fitur agen ke-i diperoleh sebagai $\mathbf{F}_i^s \leftarrow \Theta^s(X_i)$, di mana $\Theta^s(\cdot)$ adalah encoder fitur yang digunakan bersama oleh semua agen dan superskrip s mencerminkan student mode. Misalkan $\mathbf{B}_i \in \{0,1\}^{K\times K\times C}$ adalah peta BEV dari agen ke-i yang terkait dengan X_i. Komponen inti dalam fusi fitur berbasis grafik adalah *grafik kolaborasi* $\mathcal{G}_\Pi(\mathcal{V}, \mathcal{E}_\Pi)$, di mana $\mathcal{V}$ adalah himpunan simpul tetap dan $\mathcal{E}_\Pi$ adalah himpunan tepi yang dapat dilatih. Setiap simpul dalam $\mathcal{V}$ adalah agen dengan informasi pose waktu nyata; misalnya, $\boldsymbol{\xi}_i \in \mathfrak{se}(3)$ adalah pose agen ke-i dalam sistem koordinat global; dan setiap tepi dalam $\mathcal{E}_\Pi$ dapat dilatih dan memodelkan kolaborasi antara dua agen, dengan Π sebagai encoder bobot tepi, yang mencerminkan kekuatan kolaborasi yang dapat dilatih antara agen. Misalkan $\mathcal{M}_{\mathcal{G}_\Pi}(\cdot)$ adalah proses kolaborasi yang didefinisikan pada grafik kolaborasi $\mathcal{G}_\Pi$. Peta fitur dari semua agen setelah kolaborasi adalah $\{\mathbf{H}_i^s\}_{i=1}^M \leftarrow \mathcal{M}_{\mathcal{G}_\Pi}(\{\mathbf{F}_i^s\}_{i=1}^M)$. Proses ini memiliki tiga tahap: transmisi pesan neural (S1), perhatian pesan neural (S2) dan agregasi pesan neural (S3).

Pada tahap transmisi pesan neural (S1), setiap agen mengirimkan peta fitur berbasis BEV (pesan neural) ke agen lain. Pada tahap perhatian pesan neural (S2), setiap agen menerima pesan neural dari agen lain dan menentukan bobot tepi bernilai matriks ($\mathbf{W}_{j\to i}$), yang mencerminkan pentingnya pesan neural dari satu agen j ke agen i lainnya pada setiap sel individu. Karena setiap agen memiliki pose uniknya, maka grafik kolaborasi dimanfaatkan untuk mencapai transformasi fitur antar agen. Pada tahap agregasi pesan neural (S3), setiap agen menggabungkan peta fitur dari semua agen berdasarkan bobot tepi bernilai matriks yang dinormalisasi. Peta fitur yang diperbarui dari agen i adalah $\mathbf{H}_i^s = \sum_{j=1}^M \mathbf{W}_{j\to i} \odot \mathbf{F}_{j\to i}^s$, di mana $\odot$ menyatakan perkalian titik dengan penyiaran per saluran.

Metode Unggulan. V2VNet adalah yang pertama menggunakan jaringan saraf graf (GNN) yang sadar spasial untuk memodelkan komunikasi antar agen dalam persepsi kolaboratif. Selama tahap pengiriman pesan GNN, V2VNet mengompresi fitur menggunakan algoritma kompresi gambar variasional. Untuk agregasi lintas kendaraan, kerangka kerja mengkompensasi penundaan waktu untuk menetapkan keadaan awal untuk setiap node, kemudian mengubah dan mentransformasikan fitur terkompresi secara spasial dari agen tetangga ke agen ego. Operasi ini dilakukan dalam bidang pandang yang tumpang tindih. Pada tahap fusi fitur, V2VNet menggunakan operasi perataan untuk menggabungkan fitur dan memperbarui keadaan node menggunakan unit berulang gerbang konvolusional (ConvGRU).

Meskipun V2VNet menunjukkan peningkatan kinerja dengan GNN, ia menggunakan bobot kolaborasi bernilai skalar, yang tidak cukup untuk menangkap pentingnya wilayah spasial yang berbeda. Untuk mengatasi keterbatasan ini, DiscoNet memperkenalkan bobot tepi bernilai matriks untuk menangkap perhatian antar agen dengan lebih baik dalam resolusi tinggi. Selama pengiriman pesan, DiscoNet menggabungkan fitur dan menerapkan bobot tepi bernilai matriks ke setiap elemen dalam peta fitur. Selain itu, DiscoNet menggabungkan teknik fusi awal dan menengah menggunakan kerangka kerja guru-murid, yang selanjutnya meningkatkan kinerja. MP-Pose adalah kerangka kerja persepsi berbasis GNN umum lainnya. Selama tahap pengiriman pesan, MP-Pose mengkodekan hubungan spasial relatif dengan

jaringan pengkodean spasial daripada mengubah fitur secara langsung. Terinspirasi oleh Graph Attention Networks (GAT), ia selanjutnya menggunakan jaringan pengkodean perhatian silang dinamis untuk menangkap hubungan antar agen dan menggabungkan beberapa fitur seperti GAT.

Fusi Fitur Berbasis Perhatian

Selain pembelajaran grafik, mekanisme perhatian telah muncul sebagai alat yang ampuh untuk mengeksplorasi hubungan fitur. Mekanisme perhatian dapat diklasifikasikan menurut domain data menjadi perhatian saluran, perhatian spasial, dan perhatian saluran & spasial. Selama dekade terakhir, mekanisme perhatian telah memainkan peran yang semakin penting dalam visi komputer dan telah menginspirasi riset persepsi kolaboratif. Karena pemilihan fitur dan eksplorasi hubungan merupakan isu penting dalam persepsi kolaboratif perantara, beberapa karya telah memanfaatkan mekanisme perhatian untuk mengembangkan strategi fusi fitur yang lebih dinamis dan kuat. Karena fleksibilitasnya, desain berbasis perhatian telah mendominasi kolaborasi perantara.

Perumusan Masalah. Setiap agen $i \in \{1 \dots N\}$di dalam jaringan komunikasi berbagi metadata seperti pose, ekstrinsik, dan tipe agen $c_i \in \{I, V\}$satu sama lain. Memilih salah satu AV yang terhubung sebagai agen ego (e) untuk membangun sebuah graf V2X. Setiap agen $i$dapat menerima pose ego $x_e^{t_i}$pada waktu t_i. Tulang punggung PointPillar mengekstraksi peta fitur informatif $\mathbf{F}_i^{t_i} \in \mathbb{R}^{H \times W \times C}$, yang menyatakan fitur agen $i$pada waktu $t_i$dengan tinggi H, lebar W, dan saluran C. Untuk mengurangi lebar pita transmisi yang dibutuhkan, dimanfaatkan serangkaian konvolusi 1×1untuk secara progresif mengompresi peta fitur di sepanjang dimensi saluran. Fitur-fitur terkompresi dengan ukuran (H, W, C')(di mana $C' \ll C$) kemudian ditransmisikan ke agen ego (e), di mana fitur-fitur tersebut diproyeksikan kembali ke (H, W, C)menggunakan konvolusi 1×1. Fitur-fitur perantara $\mathbf{H}_i = \Gamma_\xi\left(\mathbf{F}_i^{t_i}\right) \in \mathbb{R}^{H \times W \times C}$yang diagregasi dari agen-agen yang terhubung diumpankan ke komponen utama, dengan melakukan fusi fitur antar-agen dan intra-agen secara iteratif menggunakan mekanisme atensi-diri.

***Metode Unggulan*.** Untuk menangkap interaksi di antara area spesifik dalam peta fitur, AttFusion pertama-tama menggunakan operasi self-attention pada lokasi spasial yang tepat. Secara khusus, AttFusion memperkenalkan modul fusi self-attention single-head dan mencapai keseimbangan antara kinerja dan kecepatan inferensi dibandingkan dengan metode tradisional F-Cooper dan metode berbasis grafik DiscoNet. Interaksi yang sadar spasial dalam AttFusion mirip dengan edge bobot matriks dalam DiscoNet meskipun diimplementasikan dengan alat yang berbeda. Selain pendekatan berbasis perhatian tradisional, metode berbasis Transformer telah secara signifikan memengaruhi persepsi kolaboratif. COOPERNAUT, yang terinspirasi oleh Point Transformer, memperkenalkan jaringan self-attention yang dirancang khusus untuk pemrosesan point cloud dalam pengaturan kolaboratif. Setelah menerima pesan, agen ego menggunakan blok downsampling untuk mengurangi kardinalitas himpunan titik dan blok transformator titik untuk memungkinkan pertukaran informasi lokal antar titik. Operasi ini mempertahankan invariansi permutasi pesan. Yang penting, COOPERNAUT

mengintegrasikan persepsi kolaboratif dengan keputusan kontrol, memajukan keterkaitan modul yang penting untuk mengemudi otonom.

V2X-ViT menawarkan arsitektur Transformer terpadu pertama yang secara simultan menangani interaksi kendaraan-ke-kendaraan (V2V) dan kendaraan-ke-infrastruktur (V2I). Ini memperkenalkan modul perhatian multi-agen heterogen (HMSA) untuk memodelkan hubungan yang berbeda antara agen V2V dan V2I. Selain itu, V2X-ViT menggunakan modul perhatian jendela multi-skala (MSwin) untuk menangkap interaksi spasial jarak jauh yang diperlukan untuk deteksi resolusi tinggi. Menyadari efektivitas biaya kamera RGB dibandingkan dengan LiDAR, CoBEVT menghadirkan kerangka kerja persepsi kolaboratif berbasis kamera. Modul perhatian aksial terpadu (FAX) CoBEVT memfasilitasi interaksi di antara beberapa tampilan dan agen dengan melakukan interaksi global yang jarang dan perhatian berbasis jendela lokal. Eksperimen mengkonfirmasi efektivitas CoBEVT dalam skenario multi-tampilan, multi-agen, dengan ketahanan terhadap kehilangan kamera. VIMI memperkenalkan metode Penggabungan Menengah Multi-Tampilan Kendaraan-Infrastruktur, yang berfokus pada persepsi kolaboratif berbasis kamera. Untuk mengeksplorasi korelasi antara fitur dari kendaraan dan infrastruktur pada berbagai skala, VIMI menggunakan modul perhatian silang multi-skala (MCA). Modul ini mengekstrak fitur multi-skala menggunakan CNN yang dapat berubah bentuk dan menghasilkan bobot perhatian melalui perhatian silang. Selain itu, strategi masking saluran sadar kamera (CCM) VIMI mengoreksi kesalahan kalibrasi dan meningkatkan fitur dengan menimbang ulang fitur tersebut berdasarkan parameter kamera. Studi sebelumnya terutama berfokus pada skenario kolaborasi sensor homogen.

Meskipun studi sebelumnya telah berkonsentrasi pada kolaborasi sensor homogen, eksplorasi persepsi kolaboratif dengan sensor heterogen masih kurang. HM-ViT mengatasi kesenjangan ini dengan memperkenalkan kerangka kerja persepsi kooperatif V2V heteromodal terpadu pertama. HM-ViT mencakup mekanisme perhatian grafik 3D heterogen (H3GAT) yang menggabungkan fitur pandangan mata burung (BEV) dari berbagai jenis sensor di seluruh agen. H3GAT menggunakan node dan edge yang bergantung pada tipe untuk menangkap heterogenitas sensor dan menalar tentang interaksi spasial dan hubungan antar agen. Ia juga menggabungkan perhatian berbasis jendela lokal dan perhatian berbasis grid global yang jarang untuk menangkap isyarat lokal dan global. Hasil eksperimen menunjukkan keunggulan HM-ViT dalam persepsi kolaboratif heteromodal dan homomodal.

Fungsi Kerugian

Komunikasi memperluas jangkauan persepsi agen ego, tetapi juga menghadirkan tantangan karena data bersama yang berlebihan dan tidak pasti. Untuk mengatasi masalah ini, sangat penting untuk memberikan panduan yang tepat dan mengembangkan fungsi kerugian yang disesuaikan untuk pelatihan model yang efektif. X mewakili data awan titik. Untuk deteksi objek kolaboratif deterministik, agen $\alpha \in \{1, \dots, N\}$ membuat peta *tampak atas* biner lokal $\mathcal{M} \in \{0,1\}^{W_m \times L_m \times H_m}$ dari X, di mana N adalah jumlah agen, dan W_m, L_m, dan H_m masing-masing menunjukkan lebar, panjang, dan tinggi dari peta *tampak atas* lokal tersebut. Mempertimbangkan deteksi objek kolaboratif homogen, di mana beberapa agen berbagi detektor yang sama f_θ yang terdiri dari sebuah penyandi yang dinyatakan oleh $\mathcal{E}$, dan sebuah

agregator sekaligus penyandi balik yang secara kolektif diwakili oleh $\mathcal{D}$, di mana θ menyatakan parameter-parameter dari f. $\mathcal{E}$ mengompresi peta *tampak atas* menjadi peta fitur perantara $\mathbf{F}_\alpha \in \mathbb{R}^{\frac{W_m}{K_m}\times\frac{L_m}{K_m}\times F_m} = \mathcal{E}(\mathcal{M})$, di mana K_m adalah skala penurunan sampel spasial, dan F_m adalah dimensi fitur. Setelah itu, $\mathbf{F}_\alpha$ dikirim ke agen-agen lain, dan agen α akan menggunakan $\mathcal{D}$ untuk menghasilkan sekumpulan kotak pembatas $\hat{Y} = \mathcal{D}\left(\mathbf{F}_\alpha, \{\mathbf{F}_\beta\}_{\beta\neq\alpha}\right)$.

Di dalam setiap data awan titik X, terdapat J objek. Untuk setiap objek $j \in \{1, \dots, J\}$, memprediksi I sudut dari kotak pembatas. Setiap sudut $i \in \{1, \dots, I\}$ diwakili oleh sebuah vektor berdimensi D di dalam peta *tampak atas*. Sekumpulan kotak pembatas kebenaran dasar Y diwakili sebagai $Y = \left\{c_j, \{y_{ij}\}_{i=1}^{I}\right\}_{j=1}^{J}$, di mana c adalah label klasifikasi dan $y_{ij} \in \mathbb{R}^D, \forall(i,j)$. Sekumpulan kotak pembatas yang diprediksi $\hat{Y}$ diwakili sebagai $\hat{Y} = \left\{\hat{p}_j, \{\hat{y}_{ij}, \hat{\Sigma}_{ij}\}_{i=1}^{I}\right\}_{j=1}^{J}$, di mana $\hat{p}$ adalah probabilitas klasifikasi yang diprediksi, dan setiap sudut dari kotak pembatas dimodelkan sebagai sebuah distribusi Gaussian multivariat dengan $\hat{y}_{ij} \in \mathbb{R}^D$ sebagai rata-rata yang diprediksi dan $\hat{\Sigma}_{ij}$ sebagai koordinat matriks kovarians $D \times D$ dari sudut i untuk objek j. Distribusi probabilitas dari setiap sudut kotak pembatas tersebut adalah independen. Selama pelatihan, parameter-parameter jaringan neural dari penyandi $\mathcal{E}$, agregator dan penyandi balik $\mathcal{D}$ dipelajari secara bersama-sama dengan meminimalkan kerugian deteksi $\mathcal{L}_{det}(Y, \hat{Y})$ yang mencakup sebuah kerugian klasifikasi dan sebuah kerugian regresi yang mempertimbangkan baik akurasi prediksi maupun ketidakpastian.

Metode Unggulan. Dalam lingkungan kolaboratif, agen ego sering kali menemukan data yang tumpang tindih dari agen-agen di sekitarnya. Untuk mengoptimalkan penggunaan informasi bersama ini, CRCNet, sebuah jaringan yang meningkatkan kolaborasi dengan berfokus pada komplementaritas dan mengurangi redundansi. CRCNet memiliki dua modul utama: modul peningkatan komplementaritas, yang memanfaatkan pembelajaran kontrastif untuk meningkatkan perolehan informasi, dan modul minimisasi redundansi, yang menggunakan informasi timbal balik untuk mendorong fusi fitur yang efektif. Modul-modul ini memungkinkan CRCNet untuk secara efisien menggabungkan data komplementer dari agen-agen tetangga, meningkatkan integrasi fitur secara keseluruhan. Informasi kolaboratif juga dapat mengalami ketidakpastian persepsi karena kesalahan persepsi dan kebisingan sensor, di luar sekadar redundansi. Untuk mengatasi masalah ini, metode inovatif menggunakan teknik bootstrap blok bergerak yang disesuaikan untuk menilai ketidakpastian model dan data. Selain itu, mereka mengembangkan fungsi kerugian khusus yang bertujuan untuk secara langsung memperhitungkan ketidakpastian data. Eksperimen mereka menunjukkan bahwa pendekatan ini secara efektif mengurangi ketidakpastian dan meningkatkan akurasi dalam berbagai skenario kolaboratif.

Kolaborasi Akhir

Kolaborasi akhir adalah teknik yang banyak digunakan dalam persepsi kolaboratif V2X, di mana hasil persepsi dari agen yang berbeda digabungkan pada tingkat pengambilan keputusan. Dalam pendekatan ini, keluaran sensor dari masing-masing agen diproses secara

independen, dan persepsi akhir dibentuk dengan menggabungkan hasil-hasil terpisah ini. Sementara fusi awal menawarkan keuntungan integrasi informasi kontekstual, ia membutuhkan bandwidth transmisi yang substansial. Fusi akhir, di sisi lain, memungkinkan penggabungan keluaran dari model yang berbeda tanpa memerlukan pelatihan tambahan dan menawarkan fleksibilitas untuk mengintegrasikan data dari berbagai modalitas. Namun, fusi akhir menghadirkan tantangan seperti kebutuhan akan penyelarasan yang tepat dari keluaran individu dan potensi peningkatan kompleksitas dalam proses fusi. Selain itu, ia rentan terhadap akumulasi kesalahan, yang dapat memengaruhi akurasi keseluruhan sistem persepsi kolaboratif.

Mengasumsikan terdapat dua himpunan masukan, himpunan observasi ($N = n$) dan himpunan target ($N = m$), di mana $n + m$ sama dengan jumlah total piksel di dalam gambar. Himpunan observasi adalah $O = \{(x_i, y_i)\}_{i=1}^{n}$, di mana x_i mewakili vektor fitur konteks dan y_i mewakili fitur kedalaman yang dihasilkan oleh jaringan ekstraksi fitur. Himpunan target adalah $T = \{x_i\}_{i=n+1}^{n+m}$. Sebuah fungsi pemetaan adalah $f: X \to Y$ dan misalkan P_θ adalah sebuah distribusi probabilitas pada fungsi f. Untuk $f \sim P_\theta$, kita menetapkan $\hat{y} = f(x_i)$, di mana $\hat{y}$ adalah kedalaman kebenaran dasar untuk setiap piksel. Sehingga P_θ mendefinisikan sebuah distribusi probabilitas gabungan pada $\{f(x_i)\}_{i=n+1}^{n+m}$ dan masalah penyelesaian kedalaman dirumuskan sebagai sebuah Proses Gaussian (GP) sebagai berikut:

$$P_\theta(f(T) \mid O, T) \tag{3.12}$$

di mana θ merepresentasikan seluruh parameter blok fusi induktif yang mendefinisikan distribusi P. Akibatnya, prediksi bersifat invarian terhadap permutasi dan memiliki sifat berikut, yang sesuai dengan asumsi Conditional Neural Process (CNP):

$$P_\theta(f(T) \mid O, T) = P_\theta(f(T') \mid O, T') = P_\theta(f(T) \mid O', T) \tag{3.13}$$

di mana O' dan T' *merupakan* permutasi dari O dan T. Dengan demikian, persamaan berikut dari CNP dapat diterapkan:

$$r_i = h_\theta(x_i, y_i), \quad \forall (x_i, y_i) \in O \tag{3.14}$$

$$r_j = \frac{1}{T} \sum_{i=1}^{T} r_i \tag{3.15}$$

$$\phi_i = g_\theta(x_i, r), \quad \forall x_i \in T \tag{3.16}$$

Oleh karena itu, fusi akhir tidak perlu mempertimbangkan mode pengkodean sensor yang berbeda, dan hanya perlu mencapai penyelarasan di ujung keluaran. Kolaborasi akhir biasanya mengadopsi operasi fusi pada tahap pasca-pemrosesan, yang menggabungkan keluaran persepsi multi-agen. Misalnya, kolaborasi akhir untuk deteksi objek 3D biasanya memanfaatkan metode pasca-pemrosesan seperti Non-Maximum Suppression (NMS) untuk

menghilangkan prediksi yang berlebihan dan kurang percaya diri. Namun, fusi akhir selalu menghadapi tantangan seperti ketidakselarasan spasial dan temporal.

3.5 KETAHANAN OPERASIONAL AI BERWUJUD KAWANAN

Sebagian besar studi tentang persepsi kolaboratif telah berfokus pada peningkatan efisiensi dan akurasi, sering kali mengasumsikan kondisi ideal. Namun, dalam skenario persepsi kolaboratif di dunia nyata, tantangan seperti (1) ketidakakuratan lokalisasi, (2) penundaan dan gangguan komunikasi, (3) inkonsistensi antara model atau tugas, dan (4) ancaman privasi dan keamanan, seperti yang ditunjukkan pada Gambar 3.5, dapat mengganggu efektivitas sistem ini. Masalah-masalah ini dapat menyebabkan persepsi kolaboratif berkinerja lebih rendah dibandingkan dengan persepsi individual, sehingga menimbulkan risiko keselamatan yang serius. Oleh karena itu, sangat penting untuk memastikan bahwa persepsi kolaboratif tetap kuat dan aman. Pada sub-bagian ini, kita akan membahas teknik-teknik canggih yang dikembangkan untuk mengatasi tantangan-tantangan ini.

Gambar 3.5 Tantangan dunia nyata dalam skenario persepsi kolaboratif

Kesalahan Lokalisasi

Metode persepsi kolaboratif menggunakan transformasi spasial untuk menyelaraskan dan mengintegrasikan data, fitur, atau output. Namun, masalah seperti kesalahan lokalisasi GPS dan pengukuran sensor asinkron dapat menyebabkan masalah penyelarasan, yang berdampak pada kinerja keseluruhan. Misalnya, V2VNet sangat sensitif terhadap noise pose. Pendekatan terbaru telah berfokus pada mengatasi tantangan ini dengan berbagai teknik konsistensi pose. Sebuah metode berdasarkan V2VNet diusulkan, yang meningkatkan akurasi V2VNet dengan mengintegrasikan lapisan penalaran neural yang dapat dipelajari. Metode mereka mencakup modul regresi pose yang mengoreksi transformasi yang bising dan modul konsistensi menggunakan medan acak Markov (MRF) untuk mencapai pose yang konsisten secara global di semua agen. FPV-RCNN mengatasi kesalahan lokalisasi dengan modul koreksi yang mendeteksi titik kunci pada struktur seperti tiang dan pagar. Metode ini menggunakan algoritma konsensus maksimum untuk menyelaraskan pusat agen dengan titik-titik kunci ini, memperkirakan kesalahan pose dengan lebih akurat dan mengungguli metode berbasis BEV tradisional dalam skenario dengan ketidakakuratan lokalisasi.

Metode fusi akhir, yang sering menggunakan teknik integrasi langsung, lebih rentan terhadap kesalahan lokalisasi. Untuk meningkatkan ketahanan, OptiMatch dikembangkan, sebuah sistem berdasarkan teori transportasi optimal yang menyempurnakan data tingkat objek melalui pencocokan objek yang tepat, meningkatkan akurasi bahkan dengan noise lokalisasi dan arah yang signifikan. Demikian pula, CoAlign diperkenalkan, sebuah kerangka kerja hibrida yang memperkirakan pose akurat menggunakan grafik pose agen dan objek. CoAlign mengelompokkan objek berdasarkan perkiraan ketidakpastian dan mengambil sampel pose dari kelompok-kelompok ini, mengoptimalkan konsistensi pose tanpa pengawasan, dan meningkatkan kinerja di berbagai tingkat noise.

Masalah Komunikasi

Persepsi kolaboratif di dunia nyata bergantung pada komunikasi nirkabel antar agen, yang dapat menghadapi masalah seperti latensi, gangguan, dan kehilangan informasi yang berdampak pada kolaborasi. Riset terbaru telah berfokus pada berbagai strategi untuk mengatasi tantangan ini: Untuk mengelola latensi dalam skenario fusi akhir, kerangka kerja *Time Compensation Late Fusion* (TCLF) diusulkan. TCLF memanfaatkan pelacakan dan estimasi keadaan untuk memprediksi kondisi saat ini berdasarkan frame sebelumnya. Dengan menyelaraskan prediksi dari frame yang berdekatan dan menginterpolasi posisi objek, TCLF mengurangi efek latensi dan mengintegrasikan prediksi ini dengan data kendaraan ego. Untuk fusi menengah, V2X-ViT mengatasi latensi dengan pengkodean posisi adaptif yang sadar akan penundaan (DPE) yang menyelaraskan fitur dari waktu ke waktu. Kerangka kerja ini juga menggabungkan modul HMSA dan MSwin untuk mengoreksi ketidakselarasan yang disebabkan oleh kesalahan lokalisasi dan penundaan. Eksperimen menunjukkan bahwa DPE meningkatkan kinerja meskipun terdapat penundaan waktu yang bervariasi. SyncNet dirancang untuk mengatasi latensi melalui sinkronisasi tingkat fitur.

Modul intinya, estimasi simbiotik perhatian fitur (FASE), menggunakan data waktu nyata dan historis untuk memperkirakan dan menyempurnakan fitur dan perhatian, dengan modulasi waktu yang menyesuaikan fusi fitur sesuai dengan latensi. Sebuah metode mengatasi gangguan komunikasi dengan kerangka kerja V2X-INCOP, yang memanfaatkan data historis untuk memulihkan fitur yang hilang. Kerangka kerja ini mencakup masker perhatian spasial untuk mengurangi kebisingan latar belakang dan menggunakan pembelajaran kurikulum untuk menstabilkan pelatihan. Untuk menangani kehilangan paket, yang sering disebabkan oleh hambatan atau kendaraan yang bergerak cepat, Jaringan Perbaikan Sadar LC (LCRN) diusulkan. Jaringan ini, yang terinspirasi oleh teknik pengurangan noise gambar, menggunakan arsitektur encoder-decoder untuk memulihkan fitur yang hilang selama transmisi. LCRN juga memiliki metode fusi berbasis perhatian untuk menggabungkan fitur dari berbagai agen, meningkatkan ketahanan terhadap kehilangan komunikasi.

Perbedaan Model dan Tugas

Metode persepsi kolaboratif multi-agen yang ada sering mengasumsikan bahwa semua agen menggunakan model yang identik dan dilatih untuk tugas-tugas tertentu, yang mengarah pada pembelajaran spesifik model dan spesifik tugas. Namun, dalam aplikasi dunia nyata, tidak praktis bagi semua agen untuk memiliki model yang homogen, dan pelatihan spesifik tugas

dapat membatasi skalabilitas sistem persepsi kolaboratif. Untuk mengatasi tantangan ini, beberapa studi telah mengusulkan pendekatan inovatif untuk mengurangi efek perbedaan model dan tugas. Ketika agen dilengkapi dengan model persepsi yang berbeda, metode kolaborasi yang ada dapat menghasilkan hasil fusi yang tidak dapat diandalkan karena heterogenitas model. Untuk mengatasi hal ini, kerangka kerja persepsi kolaboratif agnostik model diperkenalkan. Pendekatan ini menggunakan kalibrator offline untuk menyelaraskan skor kepercayaan agen dengan akurasi empirisnya, dengan mengakui bahwa ada distribusi kepercayaan di antara agen. Selain itu, metode ini mengusulkan modul Agregasi Promosi-Penekanan (PSA) untuk meningkatkan korelasi spasial dengan mengidentifikasi proposal promosi. Dengan kombinasi kalibrasi kepercayaan offline dan agregasi proposal online, metode kolaborasi akhir ini menunjukkan kinerja yang kuat bahkan dengan perbedaan model dan parameter.

Perbedaan model juga dapat menyebabkan kesenjangan domain antara fitur perantara, yang dapat menurunkan efektivitas kolaborasi. MPDA mengatasi masalah ini dengan mengembangkan kerangka kerja Adaptasi Domain Persepsi Multi-agen, yang pertama dari jenisnya untuk mengatasi kesenjangan domain dalam kolaborasi perantara. MPDA mencakup Pengubah Ukuran yang Dapat Dipelajari untuk menyesuaikan fitur yang diterima ke ukuran target dan menggunakan transformator lintas domain yang jarang untuk menghasilkan fitur invarian domain dengan menipu pengklasifikasi domain secara adversarial. Hasil eksperimen menunjukkan bahwa MPDA secara efektif menjembatani kesenjangan domain fitur dengan adanya perbedaan model. Pelatihan khusus tugas sering kali menghasilkan fitur yang terlalu terspesialisasi untuk tugas tertentu, sehingga membatasi generalisasinya dalam lingkungan multi-tugas. Untuk mengatasi keterbatasan ini, tugas pembelajaran mandiri baru yang disebut penyelesaian adegan multi-robot diusulkan. Tugas ini memungkinkan setiap agen untuk secara independen merekonstruksi adegan, sehingga mempelajari fitur laten yang umum dan kuat di berbagai tugas. Mereka merancang modul autoencoder spatiotemporal (STAR) untuk menyeimbangkan kinerja rekonstruksi adegan dengan efisiensi komunikasi, memungkinkan model untuk mempelajari representasi yang lebih serbaguna yang sesuai dengan berbagai tugas dalam persepsi kolaboratif.

Privasi dan Keamanan

Di luar tantangan dalam persepsi kolaboratif, privasi data dan keamanan model merupakan perhatian utama. Pelatihan kolaboratif dapat mengekspos informasi sensitif, dan model mungkin rentan terhadap serangan selama transmisi data antar agen. Bagian ini menyoroti upaya riset yang berfokus pada masalah ini, termasuk metode pelestarian privasi dan strategi pertahanan terhadap serangan. Berbagi data yang ditangkap sensor dalam persepsi kolaboratif menimbulkan risiko privasi yang signifikan, sehingga sangat penting untuk mengembangkan teknik yang menjaga privasi untuk kendaraan terhubung dan otomatis (CAV). Berbagai pendekatan telah diusulkan, seperti jaringan saraf konvolusional yang menjaga privasi (P-CNN), kerangka kerja deteksi objek awan titik yang menjaga privasi dan kooperatif tepi (SecPCV), dan kerangka kerja deteksi objek yang menjaga privasi (P2OD). Metode-metode ini menggunakan teknologi keamanan seperti pembagian rahasia aditif (ASS) dan enkripsi

kacau untuk melindungi data yang dibagikan. Namun, teknik-teknik yang menjaga privasi ini terutama telah diuji pada dataset persepsi individual, dan validasi lebih lanjut pada dataset persepsi kolaboratif diperlukan.

Persepsi kolaboratif juga bergantung pada komunikasi antar agen, yang dapat menjadi sasaran aktor jahat, sehingga membuat jaringan rentan terhadap serangan musuh. Memperkuat keamanan persepsi kolaboratif melibatkan peningkatan ketahanan terhadap serangan musuh. Hingga saat ini, hanya satu studi yang secara khusus menyelidiki serangan musuh dalam konteks ini. Metode ini mengevaluasi kinerja serangan dan pertahanan V2VNet, mengungkapkan bahwa gangguan bersama oleh penyerang menghasilkan serangan yang lebih ampuh. Di sisi pertahanan, pelatihan adversarial terbukti efektif ketika model serangan diketahui, dan riset ini juga menemukan bahwa peningkatan jumlah agen yang berkolaborasi meningkatkan kemampuan pertahanan sistem.

3.6 AI BERWUJUD KAWANAN OTONOM

Bagian ini mengeksplorasi komponen inti dari AI Otonom Berbasis Kolaborasi, dengan fokus pada kolaborasi otonom, komunikasi, dan optimasi ketahanan. Kapasitas untuk berkolaborasi adalah karakteristik utama dari AI berbasis kolaborasi, yang memungkinkan kecerdasan muncul dari tindakan kolektif banyak entitas yang berwujud. Tiga paradigma menjelaskan fenomena ini: pengorganisasian diri, di mana interaksi lokal yang sederhana menghasilkan perilaku kelompok yang kompleks tanpa pengendali pusat; negosiasi, yang memungkinkan agen untuk menggunakan penalaran eksplisit untuk bertukar proposal, menyelesaikan konflik, dan mengalokasikan tugas; dan refleksi, yang mengubah kolaborasi menjadi sistem yang belajar dari sejarahnya dengan meninjau hasil masa lalu dan menyesuaikan strategi.

Mekanisme kolaboratif ini didukung oleh komunikasi otonom, yang juga dapat dipahami melalui ketiga paradigma ini. Komunikasi berbasis pengorganisasian diri bergantung pada isyarat lingkungan tidak langsung, di mana tindakan satu agen memengaruhi persepsi agen lain, suatu proses yang dikenal sebagai stigmergi. Komunikasi berbasis negosiasi melibatkan pertukaran pesan dan balasan yang lebih langsung dan berfokus pada tujuan, memungkinkan agen untuk berbagi niat dan mencapai konsensus kelompok. Komunikasi berbasis refleksi, bentuk yang paling canggih, memungkinkan agen untuk berbagi dan memperbarui model internal yang kompleks dari proses berpikir masing-masing, mendorong model mental bersama untuk kerja tim yang mendalam.

Elemen penting terakhir adalah optimasi ketahanan otonom, kerangka kerja proaktif dan penyembuhan diri. Ini termasuk optimasi pemulihan diri, di mana swarm terus memantau kesehatannya untuk mengidentifikasi dan mengatasi kegagalan, dan optimasi konfigurasi ulang diri, yang memungkinkan swarm untuk secara dinamis mengubah struktur fisik atau logisnya untuk beradaptasi dengan kondisi baru atau kegagalan internal. Dengan mengintegrasikan paradigma ini, sistem AI swarm dapat mempertahankan ketahanan, kemampuan beradaptasi, dan peningkatan diri yang berkelanjutan di lingkungan yang dinamis.

Kolaborasi Otonom

Kemampuan untuk kolaborasi otonom telah menjadi karakteristik yang menentukan dari AI yang diwujudkan dalam bentuk swarm. Ketika kecerdasan didistribusikan di banyak entitas yang berwujud, kecerdasan tersebut tidak lagi bergantung pada keterbatasan agen tunggal. Sebaliknya, kecerdasan muncul dari proses persepsi, tindakan, dan penalaran kolektif. Kolaborasi memberikan stabilitas dan fleksibilitas dalam lingkungan yang tidak pasti, berubah, dan hanya sebagian dapat diamati. Bagian ini menguraikan tiga paradigma dominan yang menjelaskan bagaimana kawanan otonom mencapai koordinasi. Ketiga paradigma tersebut adalah kolaborasi berbasis pengaturan diri, kolaborasi berbasis negosiasi, dan kolaborasi berbasis refleksi. Setiap paradigma menyoroti mekanisme berbeda yang memungkinkan agen terdistribusi untuk mempertahankan tujuan kolektif, menyelesaikan konflik, dan beradaptasi dari waktu ke waktu.

Kolaborasi Berbasis Pengaturan Diri

Kolaborasi berbasis pengaturan diri menjelaskan bagaimana perilaku kelompok yang teratur muncul ketika agen individu berinteraksi berulang kali di bawah aturan lokal yang sederhana. Tidak diperlukan pengontrol pengawas. Sebaliknya, pola koordinasi muncul melalui penyesuaian lokal, mekanisme umpan balik, dan adaptasi terhadap kondisi eksternal. Prinsip ini memungkinkan kawanan untuk menghasilkan hasil yang tidak dapat dijelaskan hanya oleh perilaku terisolasi. Agen biasanya mengandalkan kesadaran tetangga, respons berbasis aturan, dan penguatan tindakan yang efektif. Melalui interaksi berulang, aturan-aturan ini menghasilkan pola skala besar yang tampak stabil bahkan di bawah penginderaan yang bising atau gangguan lingkungan. Ketahanan kelompok yang mengatur diri sendiri adalah salah satu alasan utama mengapa paradigma ini telah diadopsi secara luas dalam sistem kawanan yang terwujud. Jumlah agen sangat memengaruhi sifat-sifat kawanan. Menambahkan anggota baru sering kali meningkatkan kinerja keseluruhan daripada menurunkannya. Hal ini karena agen tambahan tidak hanya berkontribusi pada kemampuan mereka sendiri tetapi juga memperkuat interaksi yang mendukung koordinasi. Perilaku penskalaan seperti itu menjelaskan efektivitas pengaturan diri dalam eksplorasi terdistribusi, pemantauan area luas, dan tugas manajemen sumber daya.

Perspektif matematis tentang pengorganisasian diri sering kali bergantung pada analisis sistem nonlinier dan model konektivitas grafik. Stabilitas dijelaskan melalui loop umpan balik yang memperkuat tindakan yang berhasil dan menekan tindakan yang tidak efektif. Penundaan komunikasi atau variasi stokastik dapat mengganggu interaksi lokal, tetapi juga dapat menciptakan pola baru yang menstabilkan pada skala global. Implementasi praktis dalam kawanan yang terwujud bergantung pada pengontrol terdesentralisasi. Pengontrol ini memungkinkan agen untuk mengelola jarak, cakupan, dan penyelarasan gerakan secara otonom. Strategi konsensus atau medan potensial buatan adalah metode umum untuk menjaga kohesi global. Beberapa agen dapat berspesialisasi dalam perilaku tertentu ketika mekanisme penguatan mendukung diferensiasi, yang membantu kelompok menyeimbangkan antara eksplorasi dan eksploitasi terfokus. Tantangan di masa depan melibatkan adaptasi prinsip pengorganisasian diri terhadap kendala fisik seperti penginderaan terbatas, energi

terbatas, dan tubuh heterogen. Tidak seperti model abstrak, sistem yang terwujud membutuhkan desain yang memperhitungkan kelayakan mekanis dan batasan perangkat keras. Keterkaitan antara perwujudan dan pengorganisasian diri ini menggarisbawahi mengapa kecerdasan buatan yang diwujudkan dalam bentuk kawanan harus dipelajari sebagai bidang riset terpisah.

Kolaborasi Berbasis Negosiasi

Pengorganisasian diri menciptakan keteraturan tanpa perencanaan, tetapi kolaborasi berbasis negosiasi memperkenalkan penalaran dan penyelarasan yang eksplisit. Dalam banyak tugas, entitas otonom memiliki pengetahuan yang tidak lengkap dan kemungkinan kepentingan yang saling bertentangan. Untuk mempertahankan koordinasi yang efektif, mereka bertukar proposal, berbagi evaluasi parsial, dan menetapkan kesepakatan. Proses negosiasi ini memungkinkan kawanan untuk tetap koheren ketika mekanisme yang muncul saja tidak cukup. Negosiasi dibangun berdasarkan prinsip-prinsip teori pengambilan keputusan multi-agen dan penalaran teori permainan. Setiap peserta dimodelkan sebagai pengambil keputusan independen dengan kapasitas terbatas dan tujuan individu. Protokol negosiasi menyelaraskan tujuan-tujuan ini menjadi strategi kolektif. Protokol tersebut sering kali melibatkan komunikasi terstruktur, evaluasi utilitas, dan konvergensi iteratif menuju hasil bersama. Konflik sering terjadi dalam kawanan yang terwujud. Agen dapat menargetkan area spasial yang tumpang tindih, mencoba manuver yang tidak kompatibel, atau bersaing untuk sumber daya yang langka. Negosiasi menyediakan cara terstruktur untuk memunculkan konflik-konflik ini, bertukar prioritas, dan menentukan hasil yang menguntungkan kelompok secara keseluruhan. Proses resolusi yang disengaja ini melengkapi ketahanan yang dicapai oleh pengorganisasian diri.

Alokasi tugas merupakan aplikasi utama negosiasi. Ketika tugas berbeda dalam kompleksitas atau urgensi, mekanisme tawar-menawar memungkinkan agen untuk mendistribusikan tanggung jawab sesuai dengan kemampuan mereka. Sistem lelang dan kerangka kerja jaringan kontrak sering digunakan untuk mengimplementasikan negosiasi ini, memastikan bahwa sumber daya dan peran ditugaskan secara dinamis dan efisien. Sistem negosiasi membutuhkan protokol komunikasi dan mekanisme konvergensi. Agen harus menyeimbangkan tujuan lokal mereka sendiri dengan kendala global saat mengevaluasi pilihan. Konvergensi yang kuat memastikan stabilitas hasil bahkan ketika pesan tertunda atau hilang.Karena agen berwujud, mekanisme negosiasi harus efisien dan kuat di bawah keterbatasan perangkat keras. Komunikasi dapat tidak dapat diandalkan, bandwidth sering terbatas, dan sumber daya komputasi terbatas. Oleh karena itu, protokol negosiasi harus meminimalkan overhead sambil mempertahankan struktur yang cukup agar efektif dalam kelompok besar dan dinamis.

Kolaborasi Berbasis Refleksi

Kolaborasi berbasis refleksi memperluas kemungkinan pengaturan diri dan negosiasi dengan memperkenalkan evaluasi kolektif dan adaptasi jangka panjang. Agen tidak terbatas pada tindakan langsung. Mereka dapat meninjau interaksi masa lalu, menganalisis hasil, dan menyesuaikan strategi sesuai dengan itu. Refleksi mengubah kawanan menjadi sistem yang

dapat belajar dari sejarahnya sendiri. Refleksi melibatkan penyimpanan dan interpretasi hasil serta penyematannya dalam memori terdistribusi. Ini menutup lingkaran antara tindakan dan adaptasi dengan memungkinkan kawanan untuk merevisi aturan dan strategi berdasarkan kinerja. Hal ini menyebabkan peningkatan dari waktu ke waktu dan meningkatkan kemampuan kawanan untuk mengatasi perubahan. Pembelajaran mesin dan meta-pembelajaran sering digunakan untuk mewujudkan refleksi. Kerangka kerja pembelajaran penguatan memungkinkan kebijakan untuk diperbarui melalui akumulasi imbalan dan penalti. Meta-pembelajaran menambahkan kemampuan untuk beradaptasi dengan cepat di berbagai tugas dengan mengkodekan strategi untuk pembelajaran itu sendiri.

Memori terdistribusi sangat penting dalam sistem reflektif. Jika pengalaman tetap lokal, manfaat refleksi akan terbatas. Repositori bersama dan metode konsensus memungkinkan integrasi sejarah lokal ke dalam memori kolektif. Pengetahuan bersama ini mendukung prediksi, antisipasi perubahan lingkungan, dan penyempurnaan strategi proaktif. Refleksi juga mendukung pertimbangan keselamatan dan etika. Dengan menganalisis hasil, kawanan dapat mengenali kapan perilaku mereka bertentangan dengan batasan misi atau norma keselamatan. Oleh karena itu, refleksi menyediakan jalan untuk koreksi dan memastikan kepercayaan operasi otonom. Perwujudan memainkan peran penting dalam proses reflektif. Agen tidak hanya mengevaluasi hasil keputusan tetapi juga mengukur kinerja sensor, penggunaan energi, dan tekanan fisik. Menggabungkan dimensi-dimensi ini memastikan strategi menjadi efektif dan berkelanjutan. Tantangan riset meliputi perancangan arsitektur reflektif terdistribusi yang dapat diskalakan dengan kawanan besar, memformalkan metrik yang menangkap peningkatan dalam jangka waktu yang panjang, dan mengintegrasikan meta-pembelajaran pada tingkat kolektif. Kemajuan dalam arah ini akan memungkinkan kolaborasi berbasis refleksi menjadi mekanisme sentral dari kecerdasan kawanan yang berevolusi sendiri.

Kesimpulannya, kolaborasi otonom dalam kecerdasan buatan yang diwujudkan dalam bentuk kawanan dapat dipahami melalui tiga perspektif. Pengorganisasian diri memungkinkan koordinasi kelompok spontan muncul dari interaksi lokal. Negosiasi menyediakan pertukaran terstruktur yang menyelaraskan kepentingan dan mengalokasikan tugas. Refleksi memperkenalkan pembelajaran dari waktu ke waktu, menanamkan evaluasi dan adaptasi ke dalam kolektif. Bersama-sama, paradigma ini menjelaskan bagaimana kawanan yang diwujudkan mempertahankan ketahanan, kemampuan beradaptasi, dan peningkatan diri dalam lingkungan yang dinamis.

Komunikasi Otonom

Agar sistem kecerdasan buatan berbasis kawanan otonom dapat berfungsi dan benar-benar bermanfaat, komunikasi adalah suatu keharusan. Kemampuan agen individu untuk berbagi informasi, baik secara langsung melalui sinyal atau secara tidak langsung dengan mengubah lingkungan mereka, sangat penting bagi kelompok untuk mencapai tujuan yang tidak dapat dikelola oleh satu agen saja. Sistem komunikasi ini harus otonom, memungkinkan kawanan untuk menangani situasi yang kompleks dan selalu berubah serta masalah baru tanpa memerlukan unit kontrol pusat atau bantuan dari luar. Bagian ini membahas desain dan fungsi komunikasi otonom dengan membaginya menjadi tiga jenis utama: berdasarkan pengaturan

diri, negosiasi, dan refleksi. Setiap jenis menangani tantangan yang berbeda dan menciptakan tingkat kecerdasan kelompok yang berbeda, dari tindakan sederhana yang tidak terkoordinasi hingga strategi yang kompleks dan berorientasi pada tujuan. Menggabungkan berbagai metode komunikasi ini secara efektif adalah kunci untuk menciptakan sistem kawanan yang kuat, terukur, dan benar-benar cerdas. Gagasan di balik mekanisme ini berasal dari berbagai bidang, termasuk biologi, ilmu komputer, dan ilmu sosial.

Komunikasi Berbasis Pengorganisasian Diri

Komunikasi berbasis pengorganisasian diri adalah pendekatan terdesentralisasi, dari bawah ke atas di mana perilaku kelompok yang kompleks muncul dari interaksi lokal yang sederhana antara agen dan lingkungannya. Ini bukan tentang agen yang mengirim pesan langsung satu sama lain. Sebaliknya, ini tentang tindakan satu agen yang secara tidak langsung memengaruhi apa yang dilihat atau dirasakan oleh agen lain. Ide sentral di sini adalah stigmergi, di mana agen berkomunikasi dengan mengubah lingkungan lokal mereka, meninggalkan petunjuk yang dapat diperhatikan oleh agen lain. Perubahan ini dapat berupa jejak kimia, penanda fisik, atau bahkan pergeseran kecil dalam energi lokal. Komunikasi ini secara alami asinkron dan tidak langsung, menggunakan lingkungan sebagai saluran informasi bersama. Keuntungan besar dari model ini adalah betapa mudahnya untuk diskalakan dan betapa tangguhnya model ini. Beban komunikasi tidak meledak ketika kawanan semakin besar, karena agen hanya berinteraksi dengan tetangga terdekat mereka. Model ini juga sangat kuat terhadap kegagalan satu agen, karena tidak ada satu titik kegagalan dan perilaku keseluruhan merupakan hasil dari interaksi lokal yang tak terhitung jumlahnya.

Bagian lain dari pengorganisasian diri adalah menggunakan informasi lokal untuk mengetahui keadaan global sistem. Agen individu membuat keputusan berdasarkan apa yang mereka lihat di dekatnya, dan ketika semua keputusan ini digabungkan, pola global atau konsensus dapat terbentuk. Ini termasuk hal-hal seperti algoritma pengurutan dan agregasi terdistribusi, di mana agen bergerak atau mengubah keadaan mereka berdasarkan sifat tetangga mereka. Komunikasi di sini tidak melibatkan pertukaran data yang kompleks; ini tentang sinyal sederhana seperti menarik, menolak, atau indikator status. Misalnya, dalam pencarian makanan, agen yang menemukan sumber makanan mungkin melepaskan sinyal seperti feromon yang memengaruhi ke mana agen lain pergi selanjutnya.

Sinyal ini mungkin memudar seiring waktu, yang secara alami membuat informasi yang lebih baru menjadi prioritas yang lebih tinggi. Respons kelompok adalah perilaku yang muncul, bukan perintah yang direncanakan sebelumnya, dan efektivitasnya berasal dari pemrosesan informasi bersama yang terjadi di banyak interaksi kecil. Kelemahan utama dari metode ini adalah tidak dapat mengekspresikan banyak hal. Sinyal sederhana hanya dapat membawa sejumlah informasi yang terbatas, sehingga sulit untuk mencapai tujuan yang sangat tepat atau kompleks. Perilaku yang muncul sering kali hanya merupakan perkiraan dari solusi terbaik, dan dapat lambat untuk konvergen atau terjebak dalam keadaan suboptimal. Selain itu, karena agen tidak membagikan identitas atau keadaan mereka secara eksplisit, sulit untuk membentuk sub-kelompok yang stabil atau mengoordinasikan tugas yang membutuhkan peran atau langkah tertentu. Terlepas dari keterbatasan ini, untuk tugas-tugas seperti

eksplorasi, membuat pola terdesentralisasi, atau mencari makan dasar, pengorganisasian diri menawarkan sistem komunikasi yang sederhana, terukur, dan tangguh yang mencerminkan apa yang kita lihat dalam biologi. Riset saat ini di bidang ini berfokus pada memasukkan lebih banyak informasi ke dalam sinyal lingkungan ini dan menciptakan aturan lokal yang dapat mengarah pada bentuk kecerdasan kelompok yang lebih maju.

Komunikasi Berbasis Negosiasi

Komunikasi berbasis negosiasi adalah cara yang lebih terorganisir dan berfokus pada tujuan bagi kawanan otonom untuk bertukar informasi. Tidak seperti pengorganisasian diri, metode ini menggunakan aturan komunikasi khusus di mana agen berinteraksi langsung untuk memecahkan masalah, berbagi sumber daya, atau bekerja sama untuk mencapai tujuan bersama. Setiap agen memiliki tujuannya sendiri dan mengetahui keadaannya sendiri, dan ia juga dapat memodelkan kemungkinan keadaan dan niat agen lain. Proses komunikasi adalah serangkaian pesan dan balasan, di mana agen menyarankan ide, menanggapi dengan ide baru, dan membuat komitmen hingga mencapai kesepakatan. Hal ini dapat dijelaskan secara formal menggunakan teori permainan, di mana agen dipandang sebagai pengambil keputusan cerdas yang mencoba mendapatkan yang terbaik untuk diri mereka sendiri sambil juga membantu kelompok secara keseluruhan. Di sini, komunikasi bukan hanya tentang mengirim data; ini tentang berbagi niat, preferensi, dan mencapai konsensus kelompok.

Salah satu penggunaan utama negosiasi adalah untuk menugaskan tugas. Ketika tugas baru muncul, agen dapat menggunakan sistem lelang terdistribusi di mana mereka menawar tugas berdasarkan pekerjaan mereka saat ini, kemampuan, dan seberapa dekat mereka dengan lokasi tugas. Agen dengan tawaran terbaik mendapatkan tugas tersebut, dan informasi ini dikirim ke seluruh kawanan. Hal ini memastikan agen yang paling sesuai menangani tugas tersebut, yang membantu seluruh sistem berkinerja lebih baik. Negosiasi juga dapat digunakan untuk menghindari tabrakan dan merencanakan jalur di ruang yang ramai. Agen dapat bernegosiasi untuk hak jalan, menyarankan jalur yang berbeda untuk menghindari tabrakan dan menjaga agar semuanya berjalan lancar. Jenis komunikasi ini biasanya waktu nyata dan melibatkan banyak berbagi status langsung dan pemikiran tingkat lanjut. Seberapa baik komunikasi berbasis negosiasi bekerja bergantung pada aturan komunikasi dan kemampuan berpikir agen itu sendiri.

Pesan-pesan tersebut perlu cukup detail untuk membawa informasi yang kompleks, dan agen harus memiliki daya komputasi yang cukup untuk memproses pesan-pesan ini dan membuat keputusan yang cerdas. Tantangannya meliputi potensi hambatan komunikasi dalam kawanan besar, risiko agen bertindak buruk atau tidak mengikuti aturan, dan biaya komputasi yang tinggi untuk menjalankan algoritma negosiasi yang kompleks. Selain itu, mengandalkan seperangkat aturan tetap dapat membuat sistem kurang fleksibel dalam situasi baru atau tak terduga di mana aturannya tidak jelas. Namun untuk tugas yang membutuhkan akurasi tinggi, akuntabilitas, dan koordinasi yang direncanakan, negosiasi menawarkan cara yang ampuh untuk beralih dari perilaku kelompok sederhana ke tindakan yang disengaja dan terorganisir. Riset di bidang ini berfokus pada pembuatan aturan negosiasi yang lebih efisien, andal, dan aman yang dapat menangani lebih banyak agen dan tugas yang lebih kompleks.

Komunikasi Berbasis Refleksi

Komunikasi berbasis refleksi adalah jenis pertukaran informasi yang paling canggih dan menuntut secara kognitif dalam kawanan otonom. Ini melampaui penyampaian pesan langsung dari negosiasi dan pensinyalan tidak langsung dari pengorganisasian diri. Ini adalah sistem di mana agen dapat melihat keadaan internal mereka sendiri dan juga membangun dan memperbarui model keadaan internal dan proses berpikir agen lain. Komunikasi bukan hanya tentang berbagi data atau menyepakati sesuatu; ini tentang berbagi dan meningkatkan model internal ini. Agen dapat mengkomunikasikan seberapa yakinnya ia tentang suatu tugas, apa yang direncanakannya selanjutnya, atau mengapa ia membuat keputusan tertentu di masa lalu. Kemampuan ini memungkinkan tingkat kerja tim yang jauh lebih dalam, di mana agen dapat memprediksi apa yang dibutuhkan orang lain dan bertindak sesuai dengan itu, alih-alih hanya bereaksi terhadap apa yang terjadi.

Bagian penting dari komunikasi ini adalah gagasan tentang model mental bersama atau pemahaman kolektif tentang lingkungan dan tugas. Model bersama ini bukan hanya kumpulan fakta; ini adalah pemahaman bersama tentang bagaimana fakta-fakta tersebut saling terkait, apa tujuan keseluruhan kawanan, dan apa peran masing-masing agen. Ketika agen baru bergabung dengan kawanan atau agen lama gagal, agen lain dapat menggunakan refleksi untuk dengan cepat memperbarui model mental mereka dan menugaskan kembali tugas secara efisien. Hal ini sangat berguna dalam lingkungan yang dinamis dan tidak dapat diprediksi di mana segala sesuatunya dapat berubah dengan cepat dan membuat rencana lama menjadi tidak berguna. Jenis komunikasi ini terjadi pada tingkat yang lebih tinggi, di mana para agen memikirkan dan meningkatkan struktur kerja tim mereka.

Tantangan teknis dalam komunikasi berbasis refleksi sangat signifikan. Setiap agen membutuhkan daya komputasi yang besar untuk menjaga model internalnya tetap akurat dan mutakhir. Beban komunikasi juga bisa sangat besar, karena agen harus bertukar bukan hanya data, tetapi juga representasi kompleks dari keyakinan dan rencana mereka. Memastikan model mental bersama ini konsisten dan benar di seluruh kawanan besar merupakan tantangan besar. Jika terjadi kesalahpahaman atau penyebaran informasi yang salah, hal itu dapat menyebabkan kegagalan sistem yang besar. Namun demikian, manfaatnya juga sangat besar.

Komunikasi berbasis refleksi memungkinkan kawanan untuk menjadi fleksibel, mudah beradaptasi, dan tangguh dengan cara yang tidak dapat ditandingi oleh metode komunikasi yang lebih sederhana. Hal ini membantu dalam hal-hal seperti munculnya kepemimpinan, membentuk sub-kelompok khusus, dan menangani tugas-tugas kompleks jangka panjang yang membutuhkan pemahaman mendalam tentang masalah dan kemampuan kelompok. Riset di masa mendatang di bidang ini akan berfokus pada pengembangan cara-cara efisien untuk merepresentasikan dan mentransfer pengetahuan, serta metode untuk memastikan keandalan dan integritas model mental bersama ini dalam sistem yang terdesentralisasi.

Optimasi Ketahanan Otonom

Optimasi ketahanan otonom merupakan disiplin ilmu penting untuk menjamin utilitas dan daya tahan jangka panjang sistem swarm otonom. Pada dasarnya, paradigma ini mewakili

pergeseran dari toleransi kesalahan reaktif tradisional, dan beralih ke kerangka kerja proaktif dan penyembuhan diri. Intinya, tujuannya adalah agar swarm dapat terus mengevaluasi kesehatan dan status fungsionalnya sendiri, sehingga memungkinkan swarm untuk mengambil tindakan pencegahan dan mempertahankan kinerja bahkan ketika menghadapi potensi kegagalan. Bentuk pemeliharaan sistem tingkat lanjut ini bukan hanya tentang bertahan dari kegagalan titik tunggal; ini tentang memastikan bahwa seluruh fungsi kritis sistem tetap beroperasi meskipun terjadi degradasi yang meluas akibat malfungsi komponen hingga ketidakpastian lingkungan. Optimasi ketahanan adalah upaya multifaset, yang menggabungkan metodologi dari teori kontrol, rekayasa sistem, dan pembelajaran mesin untuk membangun entitas kolektif yang sadar diri dan mampu memperbaiki diri sendiri.

Bagian ini akan membahas konsep dasar dan teknik lanjutan dari optimasi ketahanan otonom, membedah topik menjadi dua sub-paradigma yang berbeda namun saling terkait: optimasi pemulihan diri dan optimasi konfigurasi ulang diri. Setiap pendekatan berkontribusi pada ketahanan keseluruhan suatu kawanan dengan mengatasi berbagai aspek gangguan operasional, secara kolektif memungkinkan kawanan untuk tetap efektif bahkan dalam menghadapi kesulitan yang signifikan.

Optimasi Pemulihan Diri

Tujuan utama optimasi pemulihan diri adalah untuk memungkinkan kawanan agen otonom untuk secara independen mengidentifikasi, mendiagnosis, dan mengatasi kegagalan dan konsekuensinya. Tantangan utama terletak pada arsitektur terdesentralisasi dari kawanan tersebut, yang secara inheren tidak memiliki satu titik pengawasan global. Akibatnya, agen harus bergantung pada informasi lokal dan proses kolaboratif untuk menentukan anomali dan memulai respons. Seluruh proses dimulai dengan pemantauan status berkelanjutan, di mana setiap agen menggunakan sensor internalnya untuk melacak parameter operasionalnya sendiri serta kinerja tetangganya. Setiap penyimpangan dari perilaku yang diharapkan atau pergeseran mendadak dalam tanda tangan operasional agen dapat berfungsi sebagai peringatan dini tentang potensi masalah. Misalnya, agen yang lintasan pergerakannya menyimpang dari jalur yang direncanakan, atau yang kualitas tautan datanya menurun, dapat ditandai untuk pemeriksaan lebih lanjut oleh rekan-rekannya. Penilaian kesehatan terdistribusi ini merupakan langkah pertama yang penting, menyediakan data yang diperlukan untuk diagnosis dan penyelesaian selanjutnya.

Setelah potensi masalah terdeteksi, kawanan harus menjalankan protokol diagnostik untuk menentukan sifat kesalahan. Ini sering kali merupakan tugas kolaboratif, di mana sub-kelompok agen dapat mengisolasi agen yang diduga bermasalah dan secara kolektif menjalankan rutinitas diagnostik. Rutinitas ini dapat berkisar dari pemeriksaan komunikasi dasar hingga analisis yang lebih rinci tentang variabel keadaan internal agen. Diagnosis tidak selalu mudah, karena gejala yang diamati dapat kompleks dan saling terkait. Model pembelajaran mesin tingkat lanjut, khususnya yang dilatih pada data kegagalan historis, dapat digunakan untuk membangun korelasi antara gejala dan jenis kesalahan tertentu, sehingga mempercepat proses diagnostik dan meningkatkan akurasinya. Setelah diagnosis dikonfirmasi, protokol pemulihan diaktifkan. Protokol ini bersifat adaptif dan diskalakan berdasarkan tingkat

keparahan kesalahan. Masalah kecil mungkin dapat diselesaikan dengan reboot atau kalibrasi ulang sederhana. Untuk kegagalan yang lebih parah, swarm harus mengambil tindakan yang lebih tegas.

Strategi pemulihan yang sangat efektif melibatkan penerapan redundansi secara strategis, baik dalam hal perangkat keras fisik maupun alokasi tugas. Swarm dapat dilengkapi dengan agen cadangan atau komponen duplikat yang dapat diaktifkan ketika unit utama menjadi tidak beroperasi. Bentuk redundansi fisik ini memastikan bahwa kemampuan sistem dipertahankan tanpa penurunan kinerja yang signifikan, meskipun dengan biaya peningkatan ukuran sistem dan konsumsi sumber daya. Pendekatan yang lebih canggih untuk redundansi adalah penugasan ulang tugas secara dinamis. Ketika sebuah agen didiagnosis mengalami kerusakan permanen dan dikeluarkan dari kelompok aktif, beban kerjanya harus segera didistribusikan kembali di antara agen-agen fungsional yang tersisa. Penugasan ulang ini harus cepat dan efisien untuk mencegah kegagalan misi. Kelompok agen dapat memanfaatkan lelang terdesentralisasi atau algoritma berbasis konsensus untuk mengalokasikan kembali tugas, dengan agen-agen mengajukan penawaran untuk tanggung jawab berdasarkan beban kerja dan kemampuan mereka saat ini. Bentuk redundansi operasional ini sangat fleksibel dan terukur, memungkinkan kelompok agen untuk mengalami penurunan kinerja secara bertahap daripada mengalami keruntuhan sistem total. Elemen penting dari pemulihan diri adalah kemampuan untuk belajar dari kegagalan di masa lalu.

Dengan mencatat dan menganalisis data kesalahan, kawanan dapat memperbarui model internal kesehatan sistem dan terus menyempurnakan kemampuan diagnostiknya dari waktu ke waktu. Proses pembelajaran ini dapat dikelola secara terpusat, atau secara terdistribusi. Dalam model terpusat, agen pemimpin yang ditunjuk dapat secara berkala mengumpulkan data dari semua agen lain untuk memperbarui model kesalahan global. Meskipun efektif, ini memperkenalkan satu titik kerentanan. Sebaliknya, dalam model terdesentralisasi, setiap agen dapat memelihara model lokalnya sendiri dan secara berkala bertukar pembaruan dengan tetangga, yang mengarah pada model diagnostik yang kuat yang berkembang melalui pembelajaran kolektif. Hal ini memungkinkan kawanan untuk memprediksi dan mencegah jenis kegagalan tertentu sebelum terjadi, beralih dari respons pasif ke keadaan ketahanan proaktif. Integrasi teknik pembelajaran tingkat lanjut seperti penguatan dan pembelajaran mendalam ke dalam kerangka kerja ini memiliki potensi besar untuk menciptakan sistem yang tidak hanya pulih dari kegagalan tetapi juga terus meningkatkan protokol pemulihannya melalui pengalaman operasional.

Desain mekanisme pemulihan mandiri yang efektif juga harus memperhitungkan potensi kegagalan berantai, di mana kerusakan pada satu agen memicu reaksi berantai kegagalan selanjutnya di seluruh kawanan. Ini merupakan perhatian kritis dalam sistem yang terhubung erat dengan ketergantungan antar agen yang tinggi. Protokol pemulihan yang tangguh harus mencakup langkah-langkah untuk mencegah kaskade tersebut, seperti segera mengisolasi agen yang rusak dan mengalihkan aliran komunikasi dan data di sekitar area yang terganggu. Jaringan komunikasi itu sendiri juga harus dirancang dengan redundansi dan ketahanan, memastikan bahwa informasi yang penting untuk pemulihan masih dapat

ditransmisikan meskipun sebagian jaringan terganggu. Tujuan keseluruhannya adalah untuk menciptakan sistem yang dapat menyerap guncangan dan gangguan dan tetap melanjutkan misinya, meskipun dengan kapasitas yang berpotensi berkurang. Dengan demikian, bidang optimasi pemulihan mandiri merupakan area riset masa depan yang penting untuk menciptakan sistem kawanan otonom yang andal dan tangguh dengan implikasi signifikan untuk aplikasi dalam pencarian dan penyelamatan, pemantauan lingkungan, dan eksplorasi ruang angkasa.

Optimasi Konfigurasi Ulang Mandiri

Optimasi konfigurasi ulang mandiri mengacu pada proses di mana kawanan otonom secara dinamis mengubah struktur fisik atau logisnya untuk beradaptasi dengan perubahan tujuan misi, kondisi lingkungan, atau kegagalan internal. Sementara pemulihan mandiri berfokus pada pemulihan fungsionalitas, konfigurasi ulang mandiri berkaitan dengan perubahan bentuk yang mendasar. Kemampuan ini sangat penting untuk kawanan yang benar-benar serbaguna dan cerdas. Proses konfigurasi ulang dapat melibatkan perubahan tata letak spasial kawanan, penugasan tugasnya, atau topologi komunikasi internalnya. Ini mewakili bentuk ketahanan makroskopis tingkat tinggi, yang memungkinkan kawanan untuk bergerak melampaui toleransi kesalahan sederhana ke keadaan adaptabilitas yang mendalam. Tujuan utamanya adalah agar kawanan dapat bertransisi dari satu keadaan ke keadaan lain dengan cara yang stabil dan efisien, sehingga meningkatkan kemampuannya untuk menangani lingkungan yang kompleks dan dinamis.

Salah satu bentuk konfigurasi ulang mandiri yang paling langsung adalah penyesuaian dinamis morfologi fisik kawanan. Misalnya, sekumpulan robot modular dapat terhubung atau terpisah secara fisik satu sama lain untuk membentuk struktur baru yang lebih sesuai untuk tugas tertentu. Rantai robot linier mungkin sangat efektif untuk menavigasi lorong sempit, sedangkan bidang dua dimensi yang lebih luas mungkin lebih efisien untuk menjelajahi area terbuka yang luas. Kawanan robot harus dilengkapi dengan algoritma koordinasi yang kuat untuk mengelola proses koneksi dan pemisahan, memastikan stabilitas dan fungsionalitas konfigurasi baru. Hal ini membutuhkan agen untuk memiliki pemahaman yang tepat tentang posisi dan orientasi mereka relatif terhadap tetangga mereka dan kemampuan untuk mengeksekusi manuver yang kompleks dan sinkron. Mengoptimalkan proses ini melibatkan pencarian cara paling efisien untuk beralih dari satu morfologi ke morfologi lain, meminimalkan waktu dan energi yang dihabiskan selama konversi.

Selain bentuk fisik, kawanan robot juga dapat mengkonfigurasi ulang struktur logisnya. Ini termasuk penugasan ulang peran dan tanggung jawab secara dinamis kepada masing-masing agen. Kawanan robot yang awalnya ditugaskan untuk eksplorasi, setelah menemukan target, mungkin perlu mengkonfigurasi ulang dirinya menjadi tim penyelamat. Hal ini memerlukan proses pengambilan keputusan yang terdesentralisasi di mana agen dapat secara kolaboratif memutuskan peran baru untuk setiap anggota. Proses ini dapat dimodelkan sebagai masalah optimasi global, di mana kawanan mencari alokasi tugas baru yang memaksimalkan kinerja kolektifnya dalam menghadapi tujuan baru. Konfigurasi ulang juga harus tahan terhadap kegagalan komunikasi selama transisi. Jika agen yang bertanggung jawab

atas peran penting gagal, kawanan harus dapat segera menugaskan kembali peran tersebut kepada agen lain tanpa mengganggu misi. Tingkat fleksibilitas inilah yang membedakan kawanan yang benar-benar otonom dan cerdas dari sekumpulan agen independen.

Konfigurasi ulang mandiri jaringan komunikasi kawanan merupakan aspek penting lain dari paradigma optimasi ini. Dalam lingkungan yang dinamis, tautan komunikasi tertentu dapat menjadi tidak andal atau terputus karena hambatan, gangguan, atau kegagalan agen. Kawanan harus mampu secara otonom membangun kembali komunikasi dengan membentuk tautan baru atau membuat jaringan relai untuk mengarahkan informasi di sekitar area yang terganggu. Hal ini mengharuskan agen untuk terus memantau kualitas tautan komunikasi mereka dan membagikan informasi tersebut dengan tetangga mereka. Algoritma terdesentralisasi kemudian dapat digunakan untuk menemukan topologi jaringan baru yang optimal yang memastikan konektivitas sekaligus meminimalkan latensi dan konsumsi daya. Jaringan yang dikonfigurasi ulang dapat memprioritugasan tautan komunikasi penting untuk data misi penting sementara menggunakan tautan sekunder untuk informasi yang kurang sensitif terhadap waktu. Topologi jaringan adaptif seperti itu sangat penting untuk menjaga kesadaran situasional dan koordinasi di lingkungan yang menantang.

Tantangan dalam konfigurasi ulang diri sendiri sangat signifikan. Kompleksitas ruang konfigurasi ulang dapat sangat besar, dan menemukan konfigurasi baru yang optimal secara waktu nyata merupakan tantangan komputasi yang besar. Lebih lanjut, transisi dari satu konfigurasi ke konfigurasi lain harus dieksekusi dengan presisi untuk menghindari ketidakstabilan atau kesalahan. Riset di bidang ini berfokus pada pengembangan algoritma optimasi terdesentralisasi yang dapat dengan cepat mencapai solusi yang mendekati optimal dengan overhead komunikasi minimal. Algoritma ini sering kali mengambil inspirasi dari sistem biologis yang menunjukkan bentuk konfigurasi ulang diri sendiri yang serupa, seperti automata seluler atau organisme yang dapat dikonfigurasi ulang. Tujuan utamanya adalah untuk menciptakan sistem kawanan yang tidak hanya dapat bertahan menghadapi kesulitan tetapi juga dapat memanfaatkannya sebagai katalis untuk adaptasi dan peningkatan kinerja. Kawanan yang mampu melakukan konfigurasi ulang secara otonom adalah sistem yang benar-benar tangguh dan cerdas, mampu beroperasi secara efektif dalam berbagai tugas dan lingkungan, mulai dari eksplorasi ruang angkasa hingga bantuan bencana.

3.7 SIMULATOR, DATASET, DAN METRIK

Pada bagian ini, kami memperkenalkan lingkungan simulasi pemetaan dunia nyata pertama di dunia dan kumpulan data untuk penerbangan otonom Pesawat Tanpa Awak (UAV) berkelompok. Persepsi sangat penting dalam tugas penerbangan otonom UAV untuk navigasi yang akurat dan perencanaan yang aman. Perkembangan terbaru pembelajaran mendalam membawa terobosan signifikan dalam berbagai tugas persepsi seperti deteksi objek 3D, pelacakan objek, dan segmentasi semantik. Namun, sistem visi UAV tunggal masih mengalami banyak tantangan dunia nyata, seperti oklusi dan kemampuan persepsi jarak pendek, yang dapat menyebabkan kecelakaan dahsyat dan merupakan hambatan utama dalam meningkatkan kinerja tugas ekonomi ketinggian rendah. Kekurangan tersebut terutama

disebabkan oleh bidang pandang UAV individu yang terbatas, sehingga mengakibatkan pemahaman lingkungan yang tidak lengkap. Minat yang meningkat dan kemajuan terkini dalam sistem persepsi kooperatif telah memungkinkan paradigma baru yang berpotensi mengatasi keterbatasan persepsi UAV tunggal. Dengan memanfaatkan teknologi UAV-ke-UAV (U2U), beberapa UAV (UAV) yang terhubung dan otomatis dapat berkomunikasi dan berbagi informasi sensor yang ditangkap secara bersamaan. Pada kawanan UAV yang terbang secara otonom di lingkungan terbuka yang dinamis, UAV ego kesulitan untuk mempersepsikan objek pelacakan karena halangan dedaunan dan gangguan salju. Menggabungkan fitur multimodal dari UAV di dekatnya dapat lebih lanjut mendistribusikan beberapa tugas dan mencapai fleksibilitas, ketahanan, dan jangkauan persepsi yang lebih besar, yang mengarah pada keuntungan signifikan di lingkungan yang keras dan kompleks.

Namun, terlepas dari janji yang besar, persepsi kooperatif terutama berfokus pada kendaraan dan mengabaikan tulisan masa lampau UAV, yang tetap menantang untuk memvalidasi persepsi U2U dalam skenario dunia nyata karena kurangnya dataset publik. Dataset persepsi kooperatif U2U yang ada, seperti yang ditunjukkan pada Tabel 3.1, CoPerception-UAVs dan CoPerception-UAVs+ bergantung pada simulator sumber terbuka seperti AirSim dan CARLA dan hanya mempertimbangkan 1 medan, 1 cuaca, dan 1–2 jenis sensor; mereka mengumpulkan dataset menggunakan ketinggian tetap dan mode formasi yang konsisten atau tetap. Dalam skenario dunia nyata, dibandingkan dengan mengemudi otonom, penerbangan otonom memiliki lebih banyak kebebasan, menghadapi lingkungan yang lebih kompleks, dan lebih rentan terhadap pengaruh suhu, kelembaban, dan aliran udara karena ukurannya yang lebih kecil.

Tabel 3.1 Perbandingan rinci dataset penerbangan otonom (Formasi disiplin/susunan UAV statis, FF: Formasi tetap/jalur mandiri, AF: Formasi otonom, D: Deteksi objek 3D, T: Pelacakan objek 3D, S: Segmentasi semantik).

Dataset	Tahun	Data nyata	UAV	Ketinggian	Formasi	Sampel	Medan	Cuaca	Sensor	Tugas
CoPerception -UAVs	2022	–	5	Tetap	DF, FF	4s	1	1	1	D, S
CoPerception -UAVs+	2023	–	10	Tetap	DF, FF	4s	1	1	2	D, S
U2UData	2024	Tiongkok	3	[56.6, 3000]	DF, FF, AF	0.03s	4	7	8	D, S, T

Jelas, akan ada kesenjangan domain yang jelas antara data sintetis yang ada dan data dunia nyata. Akibatnya, model yang dilatih pada dataset ini mungkin tidak dapat digeneralisasikan dengan baik ke situasi penerbangan yang realistis. Untuk lebih memajukan riset inovatif tentang persepsi kooperatif U2U, (1) kami menyajikan dataset persepsi kooperatif skala besar (U2UData) untuk penerbangan otonom swarm UAV. Data dikumpulkan oleh tiga UAV yang terbang secara otonom di U2USim, mencakup area penerbangan seluas 9 km^2, terdiri dari 315K frame LiDAR, 945K frame RGB dan kedalaman, 2,41 juta bounding box 3D beranotasi

untuk 3 kelas, dan termasuk nilai kecerahan, suhu, kelembaban, asap, dan aliran udara yang mencakup semua jalur penerbangan.

Dibandingkan dengan penerbangan pada ketinggian tetap dan mode formasi yang konsisten atau tetap, penerbangan otonom dapat menjelajahi lingkungan yang keras dan kompleks secara lebih komprehensif, memungkinkan model persepsi untuk mencapai fleksibilitas yang lebih tinggi dan ketahanan yang lebih kuat. (2) Modalitas kecerahan, suhu, kelembaban, asap, dan aliran udara dapat sangat memengaruhi kontrol penerbangan UAV. Namun, modalitas tersebut terkait erat dengan medan dan meteorologi, dan memiliki interaksi yang kuat satu sama lain, sehingga sulit untuk dikumpulkan. Kami membangun lingkungan simulasi UAV swarm pemetaan dunia nyata pertama, U2USim, yang mencakup empat medan, tujuh kondisi cuaca, dan delapan jenis sensor. (3) U2UData menggunakan Provinsi Yunnan sebagai prototipe dan menggunakan data meteorologi nyata Provinsi Yunnan yang dikumpulkan oleh Pusat Meteorologi Tiongkok untuk memetakan lingkungan simulasi berdasarkan garis bujur dan garis lintang. Hal ini memungkinkan UAV untuk merasakan dunia seperti manusia dan membuat keputusan yang optimal. (4) U2UData memperkenalkan dua tugas persepsi: deteksi objek 3D kooperatif dan pelacakan objek 3D kooperatif. (5) Kami akan membuat semua simulator, dataset, dan metrik tersedia untuk umum di seluruh dunia.

Simulator

Pada sub-bagian ini, kami membangun U2USim, lingkungan simulasi UAV yang memetakan dunia nyata pertama, dengan menggunakan Provinsi Yunnan sebagai prototipe, termasuk empat medan, tujuh kondisi cuaca, dan delapan jenis sensor. Kami menggunakan data meteorologi aktual dari Provinsi Yunnan yang dikumpulkan oleh Pusat Meteorologi Tiongkok untuk memetakan lingkungan simulasi berdasarkan garis bujur dan lintang. U2USim tidak hanya dapat mengurangi kesenjangan domain secara signifikan, tetapi UAV juga dapat menggunakannya untuk memahami dunia dan membuat keputusan optimal seperti manusia.

Simulator Umum

Simulator untuk kawanan UAV harus mampu mensimulasikan karakteristik fisik dinamis secara lebih realistis (seperti tabrakan); sensor seperti IMU, kamera, GPS, dan lidar; dan interaksi dengan ekosistem ROS. Dalam skenario dunia nyata, UAV sangat sensitif terhadap pengaruh suhu, kelembapan, dan aliran udara karena ukurannya yang kecil. Simulator UAV yang ada, seperti yang ditunjukkan pada Tabel 3.2, dapat secara visual mewujudkan digital twins di dunia nyata melalui rendering GPU, namun sulit untuk memberikan informasi modal selain modalitas visual dan LiDAR. Banyak riset yang ada secara manual menetapkan nilai berdasarkan pengalaman untuk modalitas suhu, kelembapan, kecerahan, asap, dan aliran udara. Namun, karena sensor berkaitan erat dengan medan dan meteorologi, serta memiliki interaksi yang kuat antar modalitas, model yang dilatih oleh simulator yang ada sulit diterapkan di dunia nyata.

Tabel 3.2 Perbandingan rinci simulator penerbangan otonom (H: Tinggi, M: Sedang, L: Rendah. U2USim: Berbasis data meteorologi nyata Prov. Yunnan dari Pusat Meteorologi Tiongkok sesuai bujur/lintang; memiliki sensor suhu, kelembaban, asap, & aliran udara dinamis).

Simulator	FightGear	XPlan	Jmavsim	Gazebo	AirSim	RflySim	Isaac Sim	U2USim
Sumber terbuka	–	✓	✓	✓	✓	✓	✓	✓
ROS	–	–	✓	✓	✓	✓	✓	✓
Output sensor	–	–	–	✓	✓	✓	✓	✓
Tabrakan fisik	–	–	–	✓	✓	✓	✓	✓
Rekaman	–	–	–	–	✓	✓	✓	✓
Cuaca	–	–	–	–	✓	✓	✓	✓
Data riil	–	–	–	–	–	–	Kembaran Digital	Pemetaan
Ketinggian	–	–	–	–	–	–	✓	✓
Sensor suhu	–	–	–	–	–	–	–	✓
Sensor kelembaban	–	–	–	–	–	–	–	✓
Sensor asap	–	–	–	–	–	–	–	✓
Sensor aliran udara	–	–	–	–	–	–	–	✓
Kesetiaan (Fidelity)	Rendah	Tinggi	Rendah	Sedang	Tinggi	Tinggi	Tinggi	Tinggi
Kekayaan	Rendah	Tinggi	Rendah	Sedang	Tinggi	Tinggi	Tinggi	Tinggi

Simulator kawanan UAV yang ada antara lain FlightGear, XPlane, Jmavsim, Gazebo, AirSim, RflySim, Isaac Sim, dan U2USim. Secara khusus, FlightGear bukanlah sumber terbuka. XPlane dan Jmavsim hanya dapat berinteraksi dengan ROS. Gazebo, AirSim, dan RflySim dapat berinteraksi dengan ROS, mensimulasikan tabrakan fisik, dan mengeluarkan konten sensor visual. AirSim dan RflySim juga dapat menerapkan pengendalian cuaca. Namun informasi pada simulator ini murni simulasi, dan model yang dilatih pada simulator ini sulit dijalankan di dunia nyata. Isaac Sim dan U2USim menambahkan data lingkungan nyata berdasarkan simulator sebelumnya. Isaac Sim secara visual dapat mewujudkan kembaran digital di dunia nyata melalui rendering GPU, namun sulit untuk memberikan informasi modal selain modalitas visual dan LiDAR. U2USim adalah simulator kawanan UAV pemetaan dunia nyata pertama, yang menggunakan Provinsi Yunnan sebagai prototipe, termasuk empat medan, tujuh kondisi cuaca, dan delapan jenis sensor.

Simulator Pemetaan Dunia Nyata

U2USim adalah lingkungan simulasi UAV yang memetakan dunia nyata pertama, menggunakan Provinsi Yunnan sebagai prototipe, termasuk empat medan, tujuh kondisi cuaca, dan delapan jenis sensor. Kami menggunakan data meteorologi aktual dari Provinsi Yunnan yang dikumpulkan oleh Pusat Meteorologi Tiongkok untuk memetakan lingkungan simulasi berdasarkan garis bujur dan lintang. U2USim tidak hanya dapat mengurangi kesenjangan domain secara signifikan, tetapi UAV juga dapat menggunakannya untuk

memahami dunia dan membuat keputusan optimal seperti manusia. Provinsi Yunnan terletak di barat daya Tiongkok, dengan medan tinggi di barat laut dan medan rendah di tenggara; rentang ketinggiannya adalah m. Provinsi Yunnan memiliki iklim monsun dataran tinggi subtropis dengan karakteristik iklim vertikal yang signifikan. Suhu berubah secara vertikal dengan medan, sehingga peta Provinsi Yunnan merupakan pemandangan dunia nyata yang sangat baik untuk membangun lingkungan simulasi.

U2USim menggunakan Unreal Engine (UE) 5.2 untuk membangun peta lingkungan simulasi skala kecil 3km*3km berdasarkan peta Provinsi Yunnan. U2USim mencakup empat jenis medan: pegunungan, perbukitan, dataran, dan lembah. Rentang ketinggiannya adalah m. Berdasarkan distribusi vegetasi di Yunnan, 58 jenis aset vegetasi hutan asli dibangun, dan lebih dari 15 metode superposisi digunakan untuk menggabungkannya, termasuk pertumbuhan epifit, pertumbuhan diagonal berselang-seling, dan sebagainya. Di antara itu, daun setiap tanaman akan berubah secara dinamis dengan angin, hujan, salju, dan kondisi cuaca lainnya. U2USim mengambil Provinsi Yunnan sebagai prototipe, termasuk empat medan, tujuh kondisi cuaca, dan delapan jenis sensor. Ia menggunakan data meteorologi nyata Provinsi Yunnan yang dikumpulkan oleh Pusat Meteorologi Tiongkok untuk memetakan lingkungan simulasi berdasarkan garis bujur dan garis lintang.

Sistem cuaca dirancang untuk mensimulasikan lingkungan cuaca dinamis di Provinsi Yunnan. Sistem ini terdiri dari dua bagian: kode praktis dan kinerja artistik. Untuk menciptakan efek cuaca yang mengesankan ini, digunakan 19 material dan 78 tekstur. Untuk meniru kenyataan, simulasi mencakup efek yang terjadi dari langit hingga tanah, seperti daun basah dan lantai berlumpur setelah hujan lebat di Provinsi Yunnan. Selain itu, terdapat delapan sistem partikel yang sangat disesuaikan untuk mensimulasikan hujan, salju, debu, dan kabut pada posisi tertentu dalam lingkungan simulasi. Semua efek ini terutama didasarkan pada Unreal Blueprints, yang dapat mengurangi waktu yang dibutuhkan untuk memodifikasi kode yang kompleks. Hal ini memungkinkan perancang visual untuk lebih berpartisipasi dalam proses pemrograman dan dengan mudah mengubah nilai untuk modifikasi yang realistis.

Sensor-sensor di U2USim didasarkan pada Unreal Engine, yang memungkinkan kompatibilitas lintas platform dan memberikan data yang akurat karena kemampuan perhitungan fisika yang kuat dari mesin tersebut. Sensor LiDAR terutama menggunakan deteksi sinar, dan dengan mesin yang kuat, dapat menghemat banyak waktu dalam melacak dan menghitung setiap sinar. Semua data berbasis gambar, termasuk kedalaman dan suar, dihasilkan menggunakan teknik rendering Unreal. Gambar suar dibuat melalui proses segmentasi ID, di mana objek dirender menggunakan warna suhu yang dipetakan dan dikeluarkan. Gambar kedalaman diambil selama proses pengujian kedalaman. Sensor Luminansi menggunakan kamera pada drone untuk menangkap tekstur langsung dan menghitung kecerahan rata-rata dan konsentrasi asap setiap piksel. Fungsi kecerahan didasarkan pada kurva stimulus warna mata manusia. Fungsi konsentrasi asap didasarkan pada algoritma penjumlahan gradien Laplacian. Untuk meningkatkan kredibilitas data yang diambil oleh sensor, termasuk suhu, kelembaban, dan aliran udara, data cuaca dari Provinsi Yunnan digunakan untuk dipetakan ke lingkungan simulator.

Sensor-sensor ini dipasang pada multirotor untuk menjelajahi peta simulator dan mengumpulkan data dengan interval 0,02 detik, yang dapat disesuaikan menggunakan file pengaturan JSON. Pemasangan sensor juga dapat disesuaikan; pengguna dapat mengedit file JSON untuk menyesuaikan multirotor mereka sendiri dengan memilih sensor praktis dari sembilan templat dan mendesain akurasi sensor dan tingkat pengembalian. Kami menggunakan data meteorologi nyata Provinsi Yunnan yang dikumpulkan oleh Pusat Meteorologi Tiongkok untuk memetakan lingkungan simulasi berdasarkan garis bujur dan garis lintang. Pertama, data suhu, kelembaban, tekanan, kecepatan angin, dan arah angin yang dikumpulkan oleh stasiun cuaca di Provinsi Yunnan dari tanggal 5 hingga 10 Maret 2024, yang dirilis oleh Pusat Informasi Meteorologi Nasional, diunduh. Informasi sensor nyata kemudian dipetakan ke U2USim berdasarkan informasi garis lintang dan garis bujur dari peta nyata. Di antaranya, suhu dan kelembaban adalah skalar, dan nilai yang hilang diisi dengan metode rata-rata bergerak (interval 5m). Untuk kecepatan dan arah angin, dengan asumsi bahwa kecepatan angin lintang suatu lokasi pada tanggal 5 Maret 2024 adalah u_h *m/s* dan kecepatan angin bujur adalah v_h *m/s*, maka rumus untuk kecepatan angin gabungan v_h dan arah angin θ_h adalah sebagai berikut:

$$V_h = \sqrt{u_h^2 + v_h^2} \tag{3.17}$$

$$\alpha = \arctan\left(\frac{u_h}{v_h}\right) \times \frac{180}{\pi} \tag{3.18}$$

$$\theta_h = \begin{cases} 0, & (u_h = 0,\ v_h \leq 0), \\ 90, & (u_h < 0,\ v_h = 0), \\ 270, & (u_h > 0,\ v_h = 0), \\ 180 + \alpha, & (v_h > 0), \\ 360 + \alpha, & (u_h > 0,\ v_h < 0), \\ \alpha, & (u_h < 0,\ v_h < 0) \end{cases} \tag{3.19}$$

di mana α adalah parameter sudut. Persamaan untuk kecepatan angin lintang u_h dan kecepatan angin bujur v_h pada nilai yang hilang adalah sebagai berikut:

$$u_h = \frac{u_1 \times (H_2 - H_1) + \sum_{i=2}^{n-1} u_i \times (H_{i+1} - H_{i-1}) + u_n \times (H_n - H_{n-1})}{2 \times (H_n - H_1)} \tag{3.20}$$

$$v_h = \frac{v_1 \times (H_2 - H_1) + \sum_{i=2}^{n-1} v_i \times (H_{i+1} - H_{i-1}) + v_n \times (H_n - H_{n-1})}{2 \times (H_n - H_1)} \tag{3.21}$$

di mana H_h adalah ketinggian potensial isobar (diperoleh melalui rata-rata geser titik sampel tetangga berdasarkan jarak), n adalah jumlah isobar, dan i adalah isobar pada lokasi nilai yang hilang. Kecepatan dan arah angin sintetis kemudian dihitung berdasarkan Persamaan (3.17)-(3.19).

Kumpulan Data dan Metrik

Untuk mempercepat kemajuan dalam persepsi kooperatif U2U, kami mengusulkan U2UData, kumpulan data multimodal real2sim skala besar, terbang otonom, dengan skenario cuaca yang berbeda. Kumpulan data ini dianotasi dengan cermat dengan kotak pembatas 3D untuk memfasilitasi riset tentang persepsi kooperatif swarm UAV.

Dataset Persepsi Kooperatif

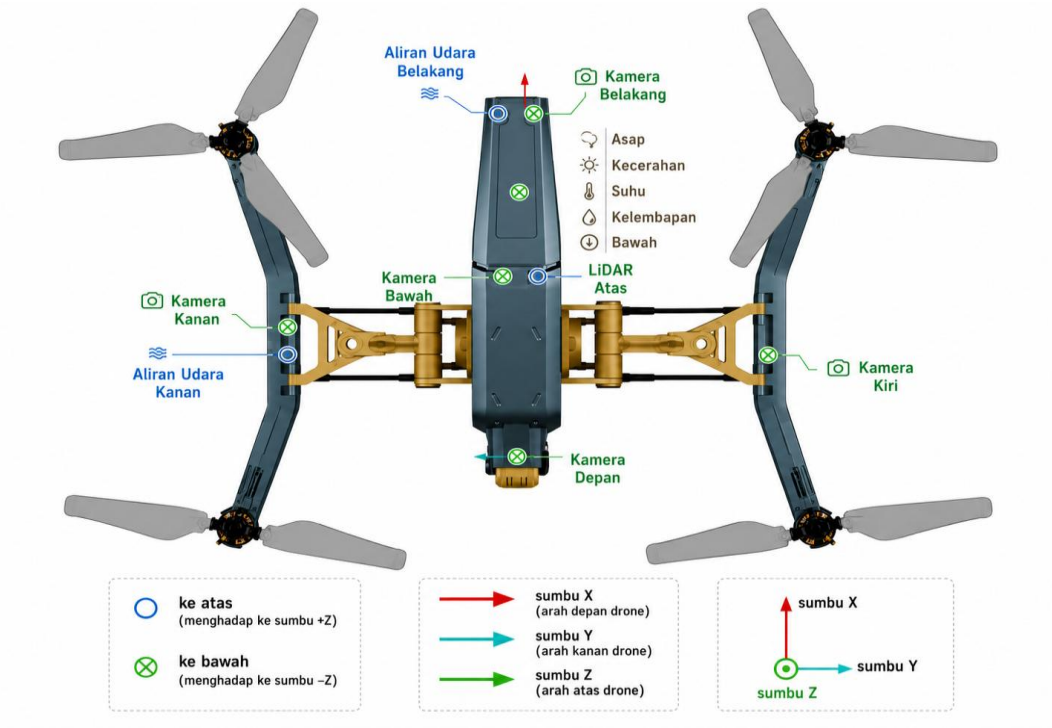
Pengumpulan kumpulan data sangat penting untuk memajukan riset algoritmik. Dataset persepsi kooperatif publik telah secara signifikan mempercepat kemajuan dalam teknologi mengemudi otonom dalam beberapa tahun terakhir. Misalnya, OPV2V memperkenalkan dataset persepsi kooperatif 3D pertama, memanfaatkan lingkungan simulasi gabungan CARLA dan OpenCDA. DAIR-V2X mempelopori dataset dunia nyata untuk deteksi kooperatif.

V2x-seq mengembangkan dataset sekuensial awal untuk persepsi dan peramalan kooperatif kendaraan-infrastruktur. V2X-Sim lebih lanjut mengeksplorasi kelayakan persepsi kendaraan-ke-segala sesuatu menggunakan data sintetis dari simulator CARLA. V2V4Real menetapkan dataset dunia nyata skala besar pertama untuk persepsi kooperatif kendaraan-ke-segala sesuatu. Namun, dataset persepsi kooperatif dalam skenario penerbangan otonom masih langka. Dataset persepsi kooperatif U2U yang ada, seperti CoPerception-UAVs dan CoPerception-UAVs+, bergantung pada simulator sumber terbuka seperti Air-Sim dan CARLA, yang memiliki medan, cuaca, dan jenis sensor yang terbatas.

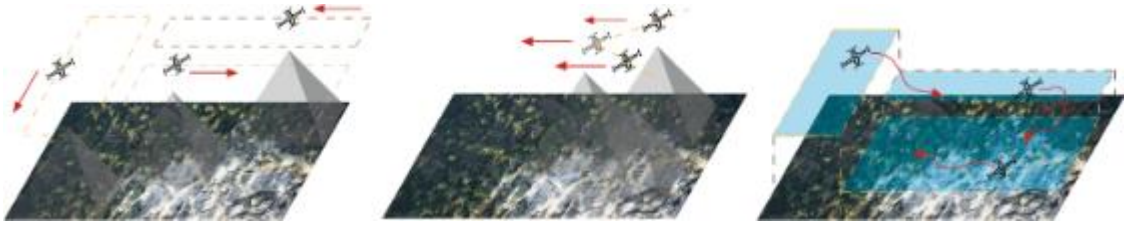
Kumpulan data ini mengumpulkan data pada ketinggian tetap dan dalam mode formasi yang konsisten atau tetap. Berbeda dengan mengemudi otonom, penerbangan otonom UAV menghadirkan kebebasan yang lebih besar, menghadapi lingkungan yang lebih kompleks, dan lebih dipengaruhi oleh cuaca alami karena ukurannya yang lebih kecil. Oleh karena itu, terdapat kesenjangan domain yang signifikan antara data sintetis yang ada dan data dunia nyata, yang berpotensi membatasi generalisasi model yang dilatih. Dalam makalah ini, kami memperkenalkan kumpulan data persepsi kooperatif skala besar pertama untuk penerbangan otonom UAV swarm.

Dataset Pemetaan Dunia Nyata

Kami mengumpulkan U2UData menggunakan tiga UAV yang terbang secara otonom di U2USim. Semua UAV dilengkapi dengan 5 kamera RGBD (depan, belakang, kiri, kanan, dan bawah), sensor 64-LiDAR (atas), 1 sensor kecerahan, suhu, kelembaban, dan asap (bawah), 2 sensor aliran udara (belakang dan kanan), dan sistem GPS/IMU. Gambar 3.6 mengilustrasikan konfigurasi tata letak sensor. Mode penerbangan UAV yang ada untuk pengumpulan data sebagian besar mengadopsi mode formasi disiplin dan formasi tetap. Seperti yang ditunjukkan pada Gambar 3.7, dalam mode formasi disiplin, kawanan UAV mempertahankan susunan yang konsisten dan relatif statis, yang memiliki ketinggian, jarak tetap, dan kecepatan yang sama. Dalam mode formasi tetap, setiap UAV bernavigasi secara independen dengan jalur tetap, memiliki ketinggian dan kecepatan yang sama.



Gambar 3.6 Konfigurasi sensor untuk platform pengumpulan data kami



Gambar 3.7 Tiga jenis formasi kawanan UAV

Namun, dalam lingkungan dunia nyata yang dinamis dan terbuka, seperti tugas pengiriman UAV, sulit bagi kawanan UAV untuk mempertahankan ketinggian, kecepatan, dan jalur tetap. Dalam U2UData, kami telah mengadopsi mode formasi otonom baru di mana setiap UAV terbang secara otonom. Dibandingkan dengan ketinggian tetap dan mode formasi disiplin atau tetap, penerbangan otonom dapat menjelajahi lingkungan yang keras dan kompleks secara lebih komprehensif, memungkinkan model persepsi untuk mencapai fleksibilitas yang lebih tinggi dan ketahanan yang lebih kuat. Untuk 7 kondisi cuaca representatif di U2USim, kami secara manual memilih 100 skenario untuk setiap kondisi cuaca, termasuk ketinggian yang berbeda dan tutupan vegetasi yang berbeda, mencakup area penerbangan seluas 9 km2. Untuk setiap skenario, kami mengumpulkan data persepsi kooperatif kawanan UAV selama 15 detik. Kami mengambil sampel bingkai gambar pada frekuensi 30Hz, yang terdiri dari total 945K bingkai RGB dan 945K bingkai kedalaman. Kami mengumpulkan 315K bingkai liDAR pada frekuensi pengambilan sampel 10Hz, termasuk 2,41M kotak pembatas 3D yang dianotasi untuk 3 kelas. 945K nilai kecerahan, 945K suhu, 945K kelembaban, 945K asap, dan 1,89M nilai aliran udara dikumpulkan melalui pemetaan dunia nyata sesuai dengan koordinat akuisisi setiap gambar.

Untuk memberi anotasi pada kotak pembatas 3D pada data LiDAR yang dikumpulkan, kami menggunakan SusTechPoint, sebuah alat pelabelan sumber terbuka yang tangguh. Kami juga secara manual memperbaiki anotasi. Terdapat total tiga kelas objek, termasuk beruang, rusa, dan serigala. Untuk setiap objek, kami memberi anotasi pada kotak pembatas 3D-nya dengan 7 derajat kebebasan, yang mencakup lokasinya (x, y, z) dan rotasinya (dinyatakan sebagai kuaternion: w, x, y, z). Lokasi (x, y, z) sesuai dengan pusat kotak pembatas. Kotak pembatas 3D ini diberi anotasi secara terpisah berdasarkan sistem koordinat global dari setiap UAV. Pendekatan ini memungkinkan data sensor dari setiap UAV untuk diperlakukan secara independen sebagai tugas deteksi agen tunggal. Kami menginisialisasi pose relatif dari kedua UAV untuk setiap frame menggunakan informasi posisi yang diberikan oleh GPS pada kedua UAV. Secara total, U2UData memiliki 315 ribu frame LiDAR, 945 ribu frame RGB dan kedalaman, 2,41 juta bounding box 3D, 1,89 juta nilai aliran udara, 945 ribu nilai kecerahan, suhu, kelembaban, dan asap. Kami secara acak membaginya menjadi set pelatihan/set validasi/set pengujian sesuai dengan rasio 0,7/0,15/0,15. Hal ini dapat sangat mempermudah kredibilitas kinerja algoritma dibandingkan dengan makalah yang berbeda.

Metrik

Berbeda dengan tugas UAV tunggal, persepsi kooperatif menghadirkan beberapa tantangan spesifik domain: Keterbatasan bandwidth. Teknologi komunikasi U2U konvensional sering beroperasi dalam batasan bandwidth yang sempit dan algoritma deteksi kooperatif perlu menyeimbangkan antara presisi dan biaya bandwidth dengan hati-hati. Kesalahan lokasi. Karena gangguan lingkungan, kesalahan bawaan dalam posisi relatif UAV yang berdekatan akan menyebabkan kesalahan pemetaan sistem koordinat. Asinkronisitas. UAV heterogen tidak dapat dihindari dan jarak antar UAV bervariasi. Frekuensi pengambilan sampel sensor dan latensi transmisi yang berbeda akan menyebabkan UAV tidak sinkron. Kami menggunakan beberapa metrik evaluasi untuk evaluasi model.

Metrik ini meliputi: AMOTA, akurasi pelacakan multiobjek rata-rata; AMOTP, presisi pelacakan multiobjek rata-rata; sAMOTA, akurasi pelacakan multiobjek rata-rata yang diskalakan, yang memastikan representasi yang lebih linier di seluruh rentang dari tugas pelacakan yang sangat menantang; MOTA, akurasi pelacakan multiobjek; MT, lintasan yang paling banyak dilacak; ML, lintasan yang paling banyak hilang.

BAB 4
KECERDASAN TERWUJUD YANG BEREVOLUSI SENDIRI

4.1 PENDAHULUAN

Bab ini memperkenalkan AI terwujud yang berevolusi sendiri, suatu bidang penting untuk mengembangkan agen otonom, seperti robot, yang dapat belajar dan beradaptasi secara langsung di dunia fisik. Kami mendefinisikan agen tersebut sebagai sistem yang berevolusi sendiri dan secara terus-menerus serta sistematis menyempurnakan model dan kebijakan internalnya melalui interaksi dinamis di dunia nyata. Bab ini menganalisis tiga teknologi utama yang memungkinkan evolusi tersebut: pembelajaran meta pembelajaran meta, yang memungkinkan adaptasi cepat terhadap tugas; pembelajaran kurikulum pembelajaran kurikulum, yang mengatur pembelajaran secara progresif dari tugas sederhana ke tugas kompleks; dan pembelajaran sepanjang hayat pembelajaran sepanjang hayat, yang memfasilitasi akumulasi pengetahuan dari waktu ke waktu tanpa mengalami kelupaan katastrofik. Selanjutnya, paradigma ini diuraikan menjadi dua komponen fundamental, yaitu model dunia yang berevolusi sendiri dan agen yang berevolusi sendiri. Komponen kedua mencakup kecerdasan yang muncul dari kognisi terwujud dan interaksi dinamis. Bab ini menyediakan peta jalan komprehensif menuju generasi baru AI yang bergeser dari pengenalan pola otonom menuju kecerdasan berbasis fisik yang berevolusi sendiri.

Munculnya model dasar berskala besar telah menandai tonggak penting dalam kecerdasan buatan, menunjukkan kemampuan yang belum pernah terjadi sebelumnya dalam pemahaman bahasa, penalaran, dan generasi. Model-model ini, yang dilatih pada kumpulan data yang luas [7–9], berfungsi sebagai mesin yang ampuh untuk generasi baru agen AI. Namun, keterbatasan mendasar tetap ada: sifat statis dari arsitektur dan basis pengetahuan yang mendasarinya. Sebagian besar sistem agen kontemporer dibangun di atas konfigurasi yang dibuat secara manual yang tetap setelah penyebaran. Paradigma desain statis ini menciptakan hambatan kritis ketika agen-agen ini dikerahkan di dunia nyata, yang pada dasarnya dinamis, tidak stasioner, dan terus berevolusi. Dalam lingkungan seperti itu, niat pengguna bergeser, persyaratan tugas berubah, dan informasi baru tersedia, sehingga agen statis menjadi semakin kurang efektif dari waktu ke waktu. Proses manual untuk mengkonfigurasi ulang dan melatih ulang sistem kompleks ini tidak hanya membutuhkan banyak tenaga dan waktu, tetapi juga pada dasarnya tidak dapat diskalakan.

Tantangan ini telah memicu pergeseran paradigma dalam pengembangan AI, beralih dari pengembangan model statis yang telah dilatih sebelumnya menuju penciptaan sistem adaptif yang mampu melakukan peningkatan diri secara otonom dan berkelanjutan. Bidang yang sedang berkembang ini, yang kami sebut AI terwujud yang berevolusi sendiri, membayangkan agen yang dapat bernalar, bertindak, dan berevolusi secara adaptif dalam waktu nyata berdasarkan interaksi dan pengalaman mereka. Bab ini memberikan gambaran komprehensif tentang bidang yang baru muncul ini, mendefinisikan konsep intinya,

mengeksplorasi teknologi pendukung utamanya, dan menguraikan komponen fundamentalnya. Transisi dari repositori pengetahuan statis ke sistem pembelajaran dinamis sepanjang hayat merupakan langkah evolusi yang diperlukan bagi AI untuk menjembatani kesenjangan antara lingkungan terkontrol dari kumpulan data pelatihan dan kompleksitas dunia nyata yang tidak dapat diprediksi. Hal ini menggeser pertanyaan sentral pengembangan AI dari sekadar apa yang diketahui model menjadi bagaimana model belajar dan beradaptasi sepanjang masa operasionalnya. Dengan mensintesis survei komprehensif terbaru, kita dapat secara formal mendefinisikan AI yang berevolusi sendiri sebagai berikut: agen AI yang berevolusi sendiri adalah sistem otonom yang terus menerus dan sistematis mengoptimalkan komponen internalnya melalui interaksi dengan lingkungannya. Tujuan utamanya adalah untuk beradaptasi dengan tugas, konteks, dan sumber daya yang berubah sambil menjaga keselamatan operasional dan meningkatkan kinerja.

Paradigma ini menjembatani kemampuan statis model dasar dengan kemampuan adaptasi berkelanjutan yang dibutuhkan oleh sistem agen sepanjang hayat, memungkinkan bentuk kecerdasan buatan yang lebih kuat dan tangguh. Untuk memastikan bahwa proses evolusi ini tetap bermanfaat dan terkendali, serangkaian prinsip panduan—analog dengan hukum robotika Asimov—telah diusulkan, yang dikenal sebagai Tiga Hukum Agen AI yang Berevolusi Sendiri. Menurut hukum-hukum ini, agen yang berevolusi sendiri harus terlebih dahulu bertahan, memastikan keselamatan dan stabilitas selama modifikasi apa pun; selanjutnya, ia harus unggul, mempertahankan atau meningkatkan kinerja tugas yang ada sesuai dengan persyaratan pertama; dan pada akhirnya, ia harus berevolusi, secara otonom mengoptimalkan komponen internalnya setelah kondisi sebelumnya terpenuhi. Secara kolektif, hukum-hukum ini menekankan bahwa pengejaran evolusi otonom harus dibatasi dengan hati-hati oleh keharusan keselamatan dan kinerja, sehingga memberikan pengaman penting untuk pengembangan teknologi yang ampuh ini secara bertanggung jawab.

Untuk menganalisis dan merancang sistem tersebut secara sistematis, bidang ini telah mengembangkan kerangka kerja konseptual yang mengabstraksikan proses inti evolusi diri. Salah satu kerangka kerja yang menonjol mengidentifikasi empat pilar yang membentuk lingkaran umpan balik yang mendorong adaptasi.

- **Masukan**. Ini termasuk data dan kondisi awal yang diberikan kepada agen, seperti spesifikasi tugas, pengamatan lingkungan, pertanyaan pengguna, dan log interaksi historis.
- **Agen**. Ini mencakup semua komponen internal agen yang tunduk pada evolusi, termasuk model dasar inti, perintah, sistem memori, pustaka alat, dan alur kerja operasional.
- **Lingkungan**. Ini mewakili konteks eksternal tempat agen beroperasi. Lingkungan berfungsi sebagai tempat pengujian kemampuan agen dan sumber sinyal umpan balik yang diperlukan untuk pembelajaran dan peningkatan.
- **Pengoptimal**. Ini adalah mekanisme algoritmik yang bertanggung jawab untuk menganalisis umpan balik dan mengimplementasikan peningkatan pada sistem agen. Contohnya termasuk optimasi berbasis gradien, algoritma pembelajaran penguatan,

dan strategi evolusioner. Kerangka kerja komplementer mengatur bidang ini berdasarkan tiga dimensi fundamental, menyediakan pendekatan terstruktur untuk mengkategorikan berbagai strategi evolusioner.

- **Apa yang Akan Dikembangkan**. Dimensi ini memeriksa komponen agen mana yang sedang dimodifikasi, seperti parameter model yang mendasarinya, struktur memori, kumpulan alat yang tersedia, atau arsitektur agen secara keseluruhan.
- **Kapan Akan Dikembangkan**. Dimensi ini mengkategorikan adaptasi berdasarkan waktunya, membedakan antara evolusi intra waktu pengujian (adaptasi cepat dalam satu contoh tugas) dan evolusi antar waktu pengujian (pembaruan yang lebih lambat dan lebih permanen antar tugas).
- **Bagaimana Cara Mengembangkan**. Dimensi ini menganalisis algoritma dan desain arsitektur spesifik yang memandu proses evolusioner, seperti metode berdasarkan imbalan skalar, umpan balik tekstual, pembelajaran imitasi, atau evolusi berbasis populasi. Secara bersama-sama, definisi dan kerangka kerja ini menyediakan landasan konseptual yang kuat untuk memahami prinsip, mekanisme, dan tujuan AI yang berevolusi sendiri dan terwujud.

Bab ini kemudian berfokus pada tiga teknologi pendukung utama: Pembelajaran Meta, yang memungkinkan agen untuk beradaptasi dengan cepat terhadap tugas-tugas baru; Pembelajaran Kurikulum, yang mengatur proses pelatihan secara progresif dari tugas-tugas sederhana hingga kompleks; dan Pembelajaran Sepanjang Hayat, yang memungkinkan agen untuk mengakumulasi pengetahuan dari waktu ke waktu sambil menghindari kelupaan yang fatal. Berdasarkan hal tersebut, bab ini selanjutnya mengeksplorasi komponen fundamental dari AI yang berevolusi sendiri dan terwujud: Model Dunia yang Berevolusi Sendiri, yang berfungsi sebagai simulator prediktif internal lingkungan untuk mendukung penalaran dan pengambilan keputusan ; dan Agen yang Berevolusi Sendiri, yang mencakup Kognisi yang Berwujud yang Berevolusi Sendiri, yang menekankan bahwa kecerdasan muncul dari keterkaitan persepsi, tindakan, dan tubuh agen, dan Interaksi Dinamis yang Berevolusi Sendiri, yang memungkinkan kolaborasi dan komunikasi yang kompleks dalam sistem multi-agen.

4.2 TEKNOLOGI UTAMA KECERDASAN TERWUJUD YANG BEREVOLUSI SENDIRI

Realisasi kecerdasan buatan terwujud yang berevolusi sendiri tidak didasarkan pada satu terobosan tunggal, melainkan pada integrasi sinergis dari beberapa paradigma pembelajaran mesin yang matang dan baru muncul. Teknologi-teknologi ini menyediakan mesin algoritmik yang memungkinkan agen untuk belajar dari pengalaman, beradaptasi dengan hal-hal baru, dan mengakumulasi pengetahuan dari waktu ke waktu. Bagian ini memberikan survei teknis terperinci tentang tiga teknologi utama: Pembelajaran Meta, yang menyediakan dasar untuk adaptasi cepat; Pembelajaran Kurikulum, yang menyusun proses pembelajaran untuk akuisisi pengetahuan yang optimal; dan Pembelajaran Sepanjang Hayat, yang memastikan retensi pengetahuan di dunia yang terus berubah.

Pembelajaran Meta

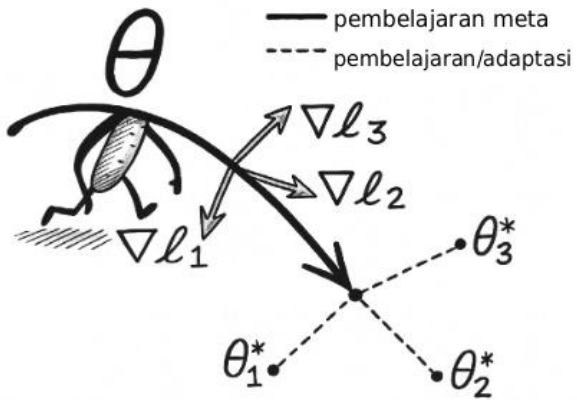
Pembelajaran meta, yang sering digambarkan sebagai “belajar untuk belajar,” adalah subbidang pembelajaran mesin yang bertujuan untuk mengembangkan model yang mampu meningkatkan algoritma pembelajarannya sendiri berdasarkan pengalaman dari berbagtugas AI. Kerangka kerja pembelajaran meta standar umumnya diuraikan menjadi dua tahap berurutan. Selama pelatihan meta, sebuah model dihadapkan pada distribusi tugas dan mempelajari strategi yang memfasilitasi transfer pengetahuan. Pada pengujian meta, model dievaluasi pada tugas-tugas yang belum pernah dilihat sebelumnya di mana ia harus beradaptasi dengan cepat hanya dengan beberapa contoh. Untuk agen yang berevolusi sendiri, kegunaan utama pembelajaran meta adalah kemampuannya untuk memberikan adaptasi yang cepat. Alih-alih mempelajari satu tugas dari awal, agen yang dilatih meta mempelajari strategi tingkat tinggi atau inisialisasi parameter yang baik yang memungkinkannya untuk menggeneralisasi dan menguastugas AI-tugas baru yang belum pernah dilihat sebelumnya dengan sangat sedikit contoh pelatihan.

Kemampuan ini sangat penting bagi agen terwujud, seperti robot, yang harus beroperasi secara efisien dan adaptif dalam berbagai lingkungan dinamis di mana pengumpulan data mahal dan memakan waktu. Proses pembelajaran meta biasanya terbagi menjadi fase pelatihan meta, di mana model belajar di seluruh distribusi tugas terkait, dan fase pengujian meta, di mana kemampuannya untuk beradaptasi dengan cepat terhadap tugas-tugas baru dievaluasi. Empat paradigma utama mendominasi tulisan masa lampau: pembelajaran meta berbasis optimasi, pembelajaran meta berbasis metrik, pembelajaran meta berbasis model, dan pembelajaran meta berbasis memori. Masing-masing mencerminkan filosofi arsitektur yang berbeda mengenai bagaimana adaptasi harus dikodekan.

Pembelajaran Meta Berbasis Optimasi

Keluarga metode ini mempelajari serangkaian parameter model awal yang secara eksplisit dioptimalkan untuk adaptasi cepat melalui beberapa langkah penurunan gradien. Karya penting di bidang ini adalah Pembelajaran Meta Agnostik-Model (MAML). MAML dibedakan oleh generalitasnya; ia kompatibel dengan model apa pun yang dapat dilatih dengan optimasi berbasis gradien, termasuk jaringan saraf dalam untuk klasifikasi, regresi, dan pembelajaran penguatan. Arsitektur MAML dicirikan oleh proses optimasi dua tingkat, seperti yang diilustrasikan pada Gambar 4.1. Putaran dalam melakukan adaptasi khusus tugas. Untuk tugas $\mathcal{T}_i$ tertentu yang diambil dari distribusi tugas $p(\mathcal{T})$, parameter model θ diperbarui menggunakan sejumlah kecil titik data (himpunan pendukung) dari tugas tersebut. Untuk satu langkah gradien, parameter khusus tugas yang telah diperbarui, $\theta_i{}'$, dihitung sebagaimana ditunjukkan pada Persamaan (4.1).

$$\theta_i{}' = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta) \tag{4.1}$$



Gambar 4.1 Diagram algoritma Pembelajaran Meta Agnostik-Model (MAML)

Di sini, $\mathcal{L}_{\mathcal{T}_i}$ adalah fungsi kerugian untuk tugas $\mathcal{T}_i$, f_θ adalah model yang diparameterisasi oleh θ, dan α adalah laju pembelajaran putaran dalam.

Putaran luar melakukan optimasi meta. Tujuannya adalah memperbaiki parameter awal θ berdasarkan kinerja parameter teradaptasi $\theta_i{}'$ pada himpunan data terpisah (himpunan kueri) untuk setiap tugas. Kerugian meta merupakan jumlah kerugian pada seluruh tugas, lalu parameter meta θ diperbarui dengan mengambil gradien kerugian tersebut. Proses ini melibatkan diferensiasi melalui langkah pembaruan putaran dalam, yaitu komputasi gradien melalui gradien, sebagaimana ditunjukkan pada Persamaan (4.2).

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}\left(f_{\theta_i{}'}\right) \tag{4.2}$$

Di sini, β adalah laju pembelajaran meta. Melalui optimasi ini, MAML menemukan inisialisasi θ yang tidak harus optimal untuk satu tugas tertentu, tetapi berada pada posisi dalam ruang parameter yang sangat peka terhadap tugas baru. Dengan demikian, model dapat memperoleh peningkatan kinerja yang substansial hanya melalui penyempurnaan minimal.

Pembelajaran Meta Berbasis Metrik

Pendekatan ini berfokus pada pembelajaran ruang penyematan fitur di mana contoh-contoh baru dapat diklasifikasikan dengan membandingkannya dengan beberapa contoh berlabel. Alih-alih mempelajari aturan pembaruan, metode ini mempelajari metrik atau fungsi jarak. Jaringan Prototipikal adalah contoh utama dari paradigma ini. Mereka beroperasi pada bias induktif bahwa ada ruang penyematan di mana titik data untuk setiap kelas mengelompok di sekitar satu titik representatif, atau prototipe. Arsitektur ini terdiri atas tiga komponen utama. Pertama, fungsi penyematan f_ϕ, yang biasanya berupa jaringan saraf dalam dengan parameter ϕ, memetakan data masukan ke ruang vektor berdimensi tinggi. Kedua, untuk setiap kelas k dalam suatu episode pembelajaran dengan sedikit contoh, vektor prototipe c_k dihitung dengan mengambil rata-rata penyematan contoh pada himpunan pendukung yang termasuk dalam kelas tersebut, sebagaimana didefinisikan pada Persamaan (4.3).

$$c_k = \frac{1}{|S_k|} \sum_{(x_i, y_i) \in S_k} f_\phi\,(x_i) \tag{4.3}$$

Di sini, S_k adalah himpunan contoh pendukung untuk kelas k. Selanjutnya, titik kueri baru x diklasifikasikan dengan menghitung distribusi probabilitas atas kelas berdasarkan fungsi normalisasi eksponensial terhadap jarak negatif ke setiap prototipe kelas, sebagaimana dinyatakan pada Persamaan (4.4). Jarak Euclidean kuadrat umumnya digunakan karena sesuai dengan asumsi distribusi Gaussian di sekitar setiap prototipe.

$$p_\phi(y = k \mid x) = \frac{\exp\left(-d\left(f_\phi(x), c_k\right)\right)}{\sum_{k'} \exp\left(-d\left(f_\phi(x), c_{k'}\right)\right)} \tag{4.4}$$

Model dilatih secara menyeluruh dengan meminimalkan log kemungkinan negatif dari kelas sebenarnya untuk titik kueri. Kesederhanaan dan efisiensi komputasi Jaringan Prototipikal menjadikannya metode yang ampuh dan efektif untuk skenario pembelajaran dengan sedikit contoh di mana adaptasi nonparametrik yang cepat diperlukan.

Pembelajaran Meta Berbasis Model

Pembelajaran meta berbasis model mengembangkan arsitektur khusus yang dirancang untuk adaptasi cepat. Alih-alih mengoptimalkan ulang parameter untuk setiap tugas baru, model ini menggunakan mekanisme internal yang mendukung generalisasi cepat. Salah satu contoh yang terkenal adalah penggunaan jaringan saraf berulang (RNN) yang memproses urutan tugas dan mengkodekan strategi spesifik tugas dalam keadaan tersembunyi. Jaringan Saraf Berbantuan Memori (MANN) memperluas ide ini dengan menggabungkan komponen memori eksternal untuk menyimpan informasi yang relevan dengan tugas. Dalam kategori ini, adaptasi dicapai bukan melalui pembaruan gradien eksplisit tetapi melalui properti arsitektur yang memungkinkan penalaran kontekstual. Ini memungkinkan adaptasi yang hampir instan setelah kapasitas representasi yang cukup diperoleh selama pelatihan meta.

Pembelajaran Meta Berbasis Memori

Pendekatan berbasis memori berfokus pada penyimpanan eksplisit pengalaman tugas masa lalu dan pengambilannya untuk adaptasi tugas baru. Metode ini dibedakan dari arsitektur berbasis model dengan penekanannya pada pengambilan memori eksternal. Jaringan Meta mempelopori kerangka kerja ini dengan mengintegrasikan bobot cepat ke dalam arsitektur saraf. Kendali Episodik Saraf dan Memori Episodik Gradien lebih lanjut mengilustrasikan bagaimana memori episodik memfasilitasi pembelajaran berkelanjutan dan mengurangi kelupaan katastrofik. Mekanisme seperti itu sangat berharga bagi agen yang beroperasi di lingkungan non-stasioner, di mana distribusi tugas berkembang seiring waktu. Dengan mempertahankan memori terstruktur dari interaksi masa lalu, model-model ini mencapai adaptasi yang kuat dan pembelajaran sepanjang hayat.

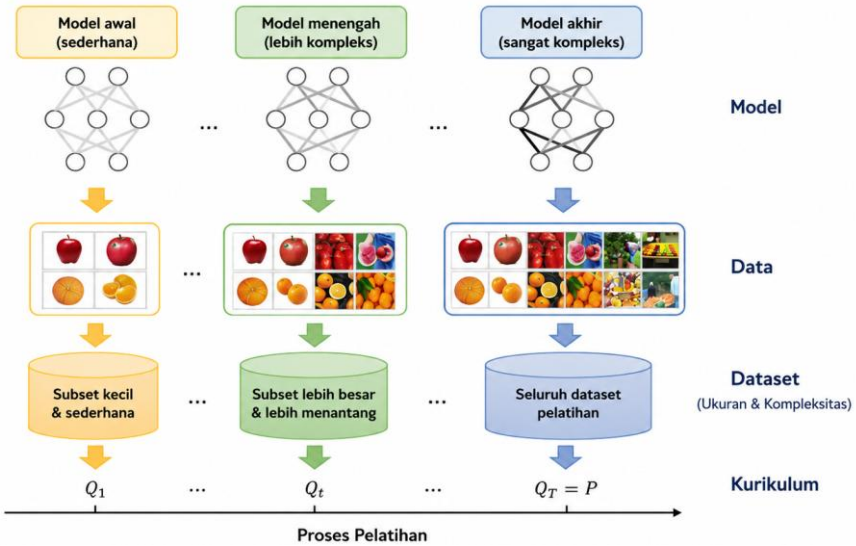
Pembelajaran Kurikulum

Pembelajaran Kurikulum (CL) adalah metodologi pelatihan yang terinspirasi oleh proses kognitif manusia dalam belajar, di mana konsep-konsep diperkenalkan dalam urutan terstruktur, biasanya berkembang dari yang sederhana ke yang lebih kompleks. Ketika diterapkan pada pembelajaran mesin, strategi ini melibatkan pengorganisasian data atau tugas pelatihan dalam urutan yang bermakna untuk memandu proses pembelajaran model. Pendekatan ini telah terbukti mempercepat konvergensi model, meningkatkan kinerja akhir, dan meningkatkan generalisasi dengan membantu pengoptimal menghindari minimum lokal yang buruk dalam lanskap kerugian. Dalam konteks pembelajaran penguatan, CL sangat berharga untuk menangani tugas-tugas dengan imbalan yang jarang atau ruang eksplorasi yang luas. Dengan memulai dengan subtugas yang lebih mudah, agen dapat memperoleh keterampilan mendasar dan membangun fungsi nilai atau representasi kebijakan yang berguna, yang kemudian dapat dimanfaatkan untuk menyelesaikan tugas target yang lebih menantang.

Karya awal oleh Bengio dan kawan kawan menetapkan dasar CL dalam pemodelan bahasa alami, menunjukkan manfaat pengurutan sampel berdasarkan kesulitan. Kumar dan kawan kawan mengusulkan Pembelajaran Mandiri (SPL), di mana model secara dinamis memilih contoh mana yang akan dipelajari berdasarkan kerangka optimasi laten. Dalam pembelajaran penguatan, Florensa dan kawan kawan memperkenalkan strategi pembangkitan tujuan otomatis, yang memungkinkan agen untuk secara progresif memperoleh keterampilan dengan menyelesaikan tugas-tugas yang semakin kompleks. Desain sistem pembelajaran kurikulum apa pun dapat dikonseptualisasikan dalam kerangka umum yang terdiri dari dua komponen penting, seperti yang ditinjau secara sistematis oleh Wang dan kawan kawan.

Pengukur Kesulitan. Komponen ini bertanggung jawab untuk mengukur kesulitan dari sampel atau tugas pelatihan tertentu. Gagasan "kesulitan" bersifat spesifik domain dan dapat didefinisikan oleh berbagai heuristik. Misalnya, dalam pemrosesan bahasa alami, panjang kalimat dapat berfungsi sebagai proksi untuk kesulitan. Dalam visi komputer, tingkat oklusi atau derau dalam gambar dapat digunakan. Atau, kesulitan dapat didefinisikan dalam kaitannya dengan keadaan model saat ini, seperti menggunakan kerugian pelatihan pada sampel sebagai indikator seberapa menantangnya bagi model.

Penjadwal Pelatihan. Komponen ini menentukan perkembangan kurikulum, menentukan bagaimana dan kapan memperkenalkan data yang lebih sulit ke model. Penjadwal dapat ditentukan sebelumnya dan statis, seperti meningkatkan panjang kalimat maksimum yang diizinkan dalam kelompok data pelatihan setiap sepuluh epok. Penjadwal yang lebih canggih bersifat dinamis dan adaptif, menyesuaikan kurikulum berdasarkan kemajuan pembelajaran model, misalnya, dengan meningkatkan kesulitan data setelah kinerja model pada himpunan bagian saat ini mencapai ambang batas tertentu (Gambar 4.2).



Gambar 4.2 Ilustrasi konsep pembelajaran kurikulum

Meskipun implementasi awal CL sering mengandalkan pakar manusia untuk mendesain kurikulum secara manual, dorongan menuju otonomi yang lebih besar telah menyebabkan pengembangan metode untuk pembuatan kurikulum otomatis. Hal ini sangat penting bagi agen yang berevolusi sendiri, yang harus mampu menyusun pengalaman belajar mereka sendiri. Beberapa strategi kunci telah muncul di domain ini.

Kerangka Kerja Pengajar-Mahasiswa Paradigma ini melibatkan dua model: model "mahasiswa" yang mempelajari tugas target, dan model "engajar" yang belajar untuk memilih atau menghasilkan subtugas yang sesuai untuk mahasiswa. Pengajar sering dimodelkan sebagai agen pembelajaran penguatan atau bandit multi-lengan, di mana "tindakannya" adalah tugas yang diusulkannya, dan "imbalannya" didasarkan pada kemajuan belajar mahasiswa. Pengajar mempelajari kebijakan untuk menyajikan tugas yang tidak terlalu mudah (tidak menawarkan informasi baru) atau terlalu sulit (tidak memberikan sinyal yang dapat dipelajari), sehingga memaksimalkan efisiensi belajar mahasiswa.

Pembelajaran Mandiri (SPL). Dalam SPL, model itu sendiri bertindak sebagai gurunya sendiri, memilih sampel mana yang akan dipelajari pada setiap tahap pelatihan. Ini biasanya dirumuskan sebagai masalah optimasi yang secara bersamaan mempelajari parameter model dan serangkaian variabel bobot laten untuk setiap sampel pelatihan. Awalnya, model memberikan bobot yang lebih tinggi pada sampel dengan kerugian pelatihan rendah (sampel "mudah"). Seiring parameter model meningkat, kepercayaannya tumbuh, dan secara bertahap memberikan bobot yang lebih tinggi pada sampel yang lebih kompleks, secara efektif menciptakan kurikulumnya sendiri.

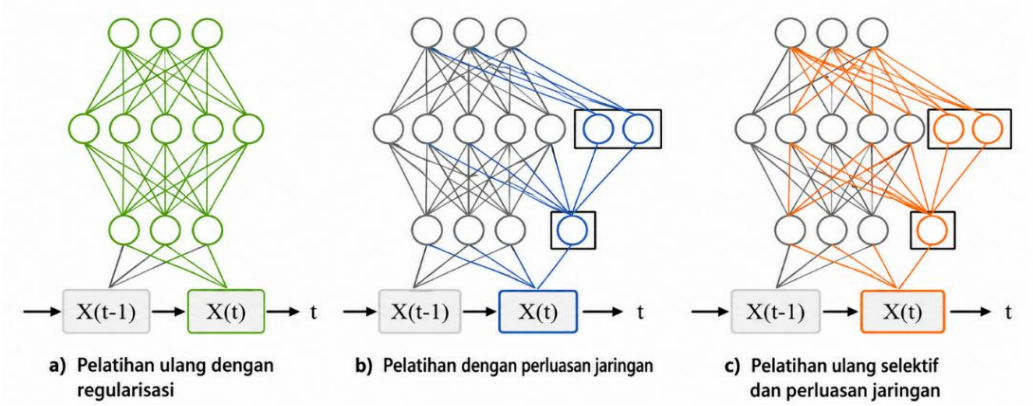
Pembangkitan Tujuan Otomatis dalam Pembelajaran Penguatan. Dalam pengaturan RL multi-tujuan, kurikulum dapat dihasilkan secara otomatis dengan memanipulasi tujuan yang disajikan kepada agen. Misalnya, untuk lengan robot yang mempelajari tugas meraih, kurikulum dapat dimulai dengan mengambil sampel tujuan yang sangat dekat dengan posisi awal lengan. Saat agen menjadi mahir dalam meraih tujuan-tujuan terdekat ini, mekanisme pembangkitan tujuan mulai mengambil sampel tujuan yang semakin jauh atau di lokasi yang

lebih menantang, seperti di balik rintangan. Strategi ini, yang sering dikombinasikan dengan teknik seperti Pemutaran Ulang Pengalaman Berbasis Tinjau Balik (HER) yang memungkinkan pembelajaran dari upaya yang gagal, secara efektif memandu proses eksplorasi dan perolehan keterampilan agen.

Komunitas pengembang telah memperluas CL ke berbagai bidang. Dalam pemrosesan bahasa alami, CL diterapkan dalam pra-pelatihan model bahasa skala besar. Dalam visi komputer, varian SPL telah digunakan untuk meningkatkan ketahanan dan efisiensi. Lebih lanjut, Matiisen dan kawan kawan mengusulkan Pembelajaran Kurikulum Guru-Siswa, di mana model guru secara adaptif merancang kurikulum untuk siswa, secara efektif menggabungkan CL dengan prinsip-prinsip pembelajaran meta. Terlepas dari keberhasilannya, beberapa masalah terbuka masih ada. Pertama, otomatisasi desain kurikulum tetap menjadi masalah yang belum terpecahkan, karena sebagian besar pendekatan bergantung pada heuristik buatan tangan atau kriteria spesifik tugas. Kedua, evaluasi efektivitas kurikulum tidak memiliki metrik standar. Terakhir, memperluas CL ke pengaturan multi-tugas dan multimoda, serta mengintegrasikannya secara mulus dengan pembelajaran meta, adalah arah yang menjanjikan tetapi masih kurang dieksplorasi.

Pembelajaran Sepanjang Hayat

Pembelajaran Sepanjang Hayat (LL), juga dikenal sebagai pembelajaran berkelanjutan, mengacu pada kemampuan suatu model untuk mempelajari banyak tugas secara berurutan tanpa melupakan pengetahuan yang telah diperoleh sebelumnya secara katastrofik. LL telah menjadi paradigma penting untuk membangun sistem AI yang beradaptasi dari waktu ke waktu dalam lingkungan yang dinamis dan terbuka. LL merupakan ciri khas kecerdasan biologis tetapi tetap menjadi tantangan besar bagi sistem buatan. Masalah utama dalam pembelajaran sepanjang hayat adalah dilema stabilitas plastisitas: suatu model harus cukup plastis (adaptif) untuk memperoleh pengetahuan dan keterampilan baru dari data baru, sekaligus cukup stabil untuk mencegah degradasi atau penghapusan pengetahuan yang telah dipelajari sebelumnya. Kegagalan untuk mempertahankan pengetahuan lama saat mempelajari tugas baru adalah fenomena yang terdokumentasi dengan baik dalam jaringan saraf yang dikenal sebagai kelupaan katastrofik atau interferensi katastrofik. Hal ini terjadi karena algoritma pelatihan standar, seperti penurunan gradien stokastik, memodifikasi bobot jaringan untuk mengoptimalkan tugas saat ini, tanpa sengaja menimpa konfigurasi bobot yang sangat penting untuk kinerja pada tugas sebelumnya. Bagi agen yang berevolusi sendiri yang ditujukan untuk penyebaran jangka panjang, mengatasi kelupaan katastrofik bukan hanya fitur yang diinginkan tetapi juga persyaratan mendasar. Untuk mengatasi tantangan ini, tiga keluarga teknik utama telah dikembangkan, masing-masing menggunakan strategi yang berbeda untuk menyeimbangkan stabilitas dan plastisitas (Gambar 4.3).



Gambar 4.3 Perbandingan konseptual dari tiga strategi pembelajaran berkelanjutan

Metode Berbasis Regularisasi. Metode ini mengurangi kelupaan dengan memperkenalkan istilah penalti tambahan ke dalam fungsi kerugian selama pelatihan pada tugas baru. Penalti ini mencegah perubahan signifikan pada parameter jaringan yang telah diidentifikasi sebagai penting untuk kinerja pada tugas sebelumnya. Algoritma penting dalam kategori ini adalah Konsolidasi Bobot Elastis (EWC). EWC menghitung bobot kepentingan per parameter berdasarkan diagonal Matriks Informasi Fisher, yang mendekati kelengkungan lanskap kerugian dan dengan demikian sensitivitas keluaran model terhadap perubahan pada parameter tersebut. Saat melatih pada tugas baru B setelah mempelajari tugas A, fungsi kerugian EWC diberikan pada Persamaan 4.5:

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i \left(\theta_i - \theta_{A,i}^*\right)^2 \tag{4.5}$$

Di sini, $\mathcal{L}_B(\boldsymbol{\theta})$ adalah kerugian untuk tugas baru, $\boldsymbol{\theta}_{A,i}^*$ adalah nilai parameter setelah pelatihan pada tugas A, F_i adalah informasi Fisher yang diestimasi untuk parameter i pada tugas A, dan λ adalah hiperparameter yang mengendalikan kompromi antara mempelajari tugas baru dan mempertahankan tugas lama. Dengan menerapkan penalti yang lebih kuat secara selektif pada bobot yang lebih penting, EWC membentuk kendala "elastis" yang memungkinkan bobot kurang penting berubah dengan lebih bebas.

Metode Berbasis Pemutaran Ulang. Metode berbasis pemutaran ulang (atau berbasis memori) mengatasi kelupaan dengan menyimpan sebagian kecil data dari tugas-tugas sebelumnya dalam penyangga memori eksplisit. Selama pelatihan pada tugas baru, sampel dari penyangga ini "diputar ulang" bersamaan dengan data baru, secara efektif menciptakan kelompok data campuran yang mengingatkan model akan pengetahuan sebelumnya. Latihan langsung dari pengalaman masa lalu ini adalah strategi yang sangat efektif, meskipun secara konseptual sederhana, untuk mencegah kelupaan katastrofik. Tantangan yang signifikan adalah biaya memori dan potensi masalah privasi yang terkait dengan penyimpanan data mentah. Untuk mengatasi hal ini, Pemutaran Ulang Generatif telah diusulkan. Dalam pendekatan ini, alih-alih menyimpan sampel data, model generatif (seperti Pengode Otomatis Variasional atau Jaringan Adversarial Generatif) dilatih pada data setiap tugas. Saat

mempelajari tugas baru, model generatif ini digunakan untuk menghasilkan "sampel semu" dari tugas-tugas sebelumnya, yang kemudian digunakan untuk pemutaran ulang, sehingga menghindari kebutuhan untuk menyimpan data asli.

Metode Arsitektur. Metode ini mencegah interferensi antar tugas dengan memodifikasi arsitektur model secara dinamis. Ide intinya adalah mengalokasikan serangkaian parameter yang berbeda untuk tugas yang berbeda, sehingga mengisolasi basis pengetahuan masing-masing. Contoh mendasar adalah Jaringan Saraf Progresif. Ketika model dihadapkan dengan tugas baru, "kolom" jaringan saraf baru ditambahkan ke arsitektur. Parameter dari semua kolom yang telah dilatih sebelumnya dibekukan, sehingga tidak mungkin terjadi kelupaan. Setiap lapisan dalam kolom baru menerima masukan tidak hanya dari lapisan sebelumnya dalam kolomnya sendiri, tetapi juga, melalui koneksi lateral, dari lapisan sebelumnya dari semua kolom yang telah dilatih sebelumnya. Hal ini memungkinkan model untuk memanfaatkan fitur yang dipelajari dari tugas-tugas sebelumnya (transfer maju positif) sambil memastikan bahwa pengetahuan lama tetap terjaga dengan sempurna. Kelemahan utama dari pendekatan ini adalah ukuran model tumbuh secara linier dengan jumlah tugas, yang dapat menjadi tidak berkelanjutan untuk agen pembelajaran sepanjang hayat yang sesungguhnya.

Bidang ini telah berkembang pesat menuju pembelajaran berkelanjutan yang tidak bergantung pada tugas, di mana batasan tugas tidak diberikan secara eksplisit, membuat pembelajaran lebih realistis tetapi juga lebih menantang. Pada saat yang sama, para ahli menekankan efisiensi: menyeimbangkan biaya komputasi dan kendala memori sangat penting untuk meningkatkan skala LL ke aplikasi dunia nyata. LL juga semakin dikombinasikan dengan pembelajaran meta dan pembelajaran kurikulum, membentuk sistem adaptif yang dapat mengingat tugas-tugas sebelumnya dan melakukan generalisasi dengan cepat ke tugas-tugas baru. Tantangan utama masih belum terselesaikan. Pertama, pencegahan kelupaan katastrofik tanpa memori atau pertumbuhan parameter yang tidak terbatas masih merupakan masalah riset yang terbuka. Kedua, merancang tolok ukur yang mencerminkan skenario realistis sepanjang hayat, alih-alih tugas yang dipisahkan secara artifisial, sangat penting untuk mengevaluasi kemajuan. Terakhir, mengembangkan kerangka kerja terpadu yang mengintegrasikan LL dengan paradigma lain, seperti pembelajaran meta dan CL, merupakan arah masa depan yang menjanjikan.

Ketiga paradigma yang dibahas di atas membahas aspek adaptabilitas dan ketahanan yang berbeda. Perbandingan yang jelas tentang tujuan, mekanisme, dan peran mereka dapat menyoroti bagaimana mereka saling melengkapi dalam membentuk sistem kecerdasan buatan yang berevolusi sendiri. Tabel 4.1 memberikan ringkasan terstruktur dari paradigma-paradigma ini. Analisis komparatif dalam Tabel 4.1 menunjukkan bahwa ketiga paradigma tersebut membahas dimensi kecerdasan adaptif yang berbeda tetapi saling terkait. Pembelajaran Meta memastikan efisiensi dalam memperoleh tugas-tugas baru, Pembelajaran Kurikulum menekankan pengorganisasian progresif perolehan pengetahuan, dan Pembelajaran Sepanjang Hayat menjaga stabilitas di berbagai rentang waktu yang panjang. Bersama-sama, pendekatan-pendekatan ini merupakan perangkat pelengkap untuk

mengembangkan agen yang mampu berevolusi secara berkelanjutan dan mandiri di lingkungan yang kompleks.

4.3 KOMPONEN UTAMA KECERDASAN TERWUJUD YANG BEREVOLUSI SENDIRI

Setelah menetapkan teknologi dasar, bagian ini menguraikan konsep kecerdasan buatan yang berevolusi sendiri menjadi tiga pilar inti yang saling bergantung. Otonomi sejati pada agen buatan tidak hanya membutuhkan kepemilikan kemampuan canggih tetapi juga kapasitas untuk secara aktif menyempurnakan dan meningkatkannya selama interaksi sepanjang hayat. Evolusi ini didasarkan pada siklus pembelajaran, adaptasi, dan interaksi yang berkelanjutan. Pilar pertama adalah model dunia yang berevolusi sendiri, pemahaman prediktif internal agen tentang lingkungannya yang harus terus beradaptasi untuk mempertahankan kesetiaan di dunia yang tidak stasioner. Yang kedua adalah kognisi yang berevolusi sendiri, yang mendasarkan pemahaman ini pada interaksi fisik, memungkinkan kecerdasan untuk muncul dan berkembang dari interaksi timbal balik antara tubuh agen, persepsinya, dan lingkungannya. Pilar ketiga dan terakhir, interaksi dinamis yang berevolusi sendiri, memperluas prinsip-prinsip ini ke sistem multi-agen, di mana kecerdasan kolektif dan struktur sosial muncul melalui koordinasi, komunikasi, dan adaptasi timbal balik yang dipelajari. Bersama-sama, komponen-komponen ini membentuk cetak biru bagi agen yang tidak hanya menjalankan perilaku yang telah diprogram sebelumnya, tetapi juga tumbuh, belajar, dan berevolusi menjadi entitas yang lebih mampu dan cerdas.

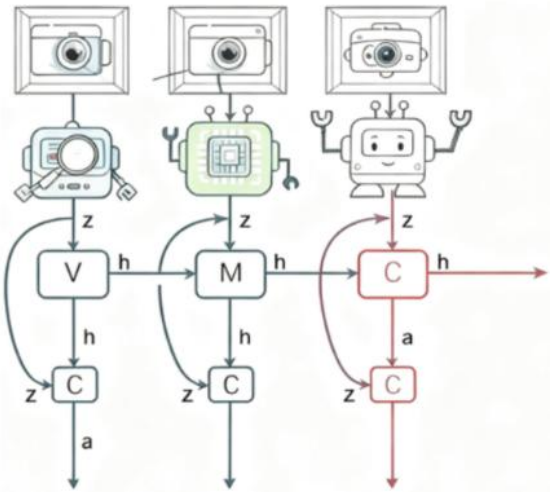
Tabel 4.1 Analisis komparatif teknologi pendukung utama untuk AI terwujud yang berevolusi sendiri

Teknologi	Tujuan utama	Mekanisme inti	Peran dalam AI yang berevolusi sendiri
Pembelajaran meta	Adaptasi cepat terhadap tugas-tugas baru dengan data minimal	Mempelajari proses optimasi atau ruang metrik di seluruh distribusi tugas	Memungkinkan agen untuk belajar belajar, memberikan kemampuan mendasar untuk menyesuaikan keterampilan dan strateginya secara efisien dalam situasi baru
Pembelajaran kurikulum	Meningkatkan konvergensi dan generalisasi dengan menyusun pengalaman belajar	Mengorganisir data atau tugas pelatihan dalam urutan yang bermakna, biasanya dari yang mudah ke yang sulit	Memungkinkan agen untuk secara mandiri menyusun jalur pembelajarannya sendiri, memperoleh keterampilan dasar pada masalah yang lebih sederhana sebelum menangani masalah yang lebih kompleks, sehingga mengoptimalkan lintasan perkembangannya
Pembelajaran sepanjang hayat	Mengumpulkan pengetahuan dari waktu ke waktu tanpa melupakan secara tiba-tiba.	Mengurangi interferensi antar tugas yang dipelajari secara berurutan melalui regularisasi, pemutaran ulang, atau modifikasi arsitektur.	Memberikan stabilitas yang diperlukan untuk otonomi jangka panjang, memastikan bahwa seiring agen berevolusi dan memperoleh pengetahuan baru, ia tidak kehilangan kemampuan yang telah dikuasai sebelumnya.

Model Dunia yang Berevolusi Sendiri

Model dunia adalah representasi internal yang dipelajari agen untuk merangkum dinamika lingkungannya. Secara fungsional, ini adalah jaringan saraf generatif yang mempelajari representasi spasial dan temporal yang terkompresi, berfungsi sebagai mesin prediksi atau simulator internal. Kemampuan ini memungkinkan agen untuk memprediksi konsekuensi masa depan dari tindakannya tanpa mengeksekusinya di dunia nyata. Dengan menanyakan model dunianya, agen dapat "membayangkan" atau "memimpikan" lintasan masa depan, memungkinkan perencanaan dan pengambilan keputusan yang canggih. Paradigma ini merupakan dasar dari pembelajaran penguatan berbasis model, karena telah ditunjukkan bahwa agen dapat dilatih sepenuhnya dalam lingkungan simulasinya sendiri, dengan kebijakan yang dihasilkan berhasil ditransfer kembali ke lingkungan sebenarnya. Bagi agen yang beroperasi di dunia yang dinamis dan tidak stasioner, model dunia statis tidak cukup. Model itu sendiri harus berevolusi untuk mempertahankan kesetiaannya dengan realitas yang berubah.

Hambatan utama dalam pengembangan model dunia yang berevolusi sendiri adalah kelupaan katastrofik. Fenomena ini terjadi ketika jaringan saraf, setelah mempelajari tugas baru atau beradaptasi dengan dinamika lingkungan baru, menimpa bobot sinaptik yang mengkodekan pengetahuan yang diperoleh sebelumnya. Model dunia yang berevolusi sendiri oleh karena itu membutuhkan mekanisme khusus yang dirancang secara eksplisit untuk mengelola pertukaran antara stabilitas dan plastisitas (Gambar 4.4).



Gambar 4.4 Diagram konseptual arsitektur model dunia

Pembelajaran Sepanjang Hayat melalui Memori dan Pemutaran Ulang. Untuk mengatasi kelupaan katastrofik, model dunia harus dilengkapi dengan mekanisme untuk pembelajaran sepanjang hayat atau berkelanjutan. Salah satu strategi langsung adalah penggunaan penyangga pemutaran ulang pengalaman yang bertahan di seluruh tugas. Saat agen belajar, ia menyimpan pengalaman dalam penyangga bersama. Ketika model dunia dilatih, ia mengambil sampel tidak hanya dari pengalaman terkini tetapi juga dari penyangga historis ini, secara efektif melatih tugas-tugas masa lalu sambil mempelajari tugas-tugas baru.

Alternatif yang lebih canggih adalah Deep Pemutaran Ulang Generatif, di mana agen melatih model generatif untuk menangkap distribusi data dari tugas-tugas sebelumnya. Ketika tugas baru diperkenalkan, generator ini menciptakan data semu sintetis yang meniru pengalaman masa lalu. Sampel yang dihasilkan ini kemudian diselingi dengan data nyata dari tugas baru untuk melatih model utama agen, memungkinkannya untuk mempelajari tugas baru tanpa melupakan tugas-tugas lama.

Pemodelan Struktur Lingkungan Dinamis. Paradigma yang lebih maju untuk evolusi model dunia berfokus pada pemahaman struktur mendasar dari perubahan lingkungan. Hal ini dapat dicapai dengan mengkonseptualisasikan lingkungan sebagai grafik dinamis, di mana entitas adalah simpul dan hubungannya adalah tepi. Perubahan dalam lingkungan dapat dimodelkan secara formal sebagai pergeseran distribusi spasial-temporal dalam struktur grafik yang mendasarinya. Jaringan Saraf Grafik Dinamis dirancang untuk belajar pada struktur yang berevolusi tersebut. Kerangka kerja seperti Pembelajaran Grafik Dinamis Sadar Lingkungan (EAGLE) secara eksplisit memodelkan dan menyimpulkan variabel lingkungan laten yang mengatur evolusi grafik, memungkinkan model untuk mengidentifikasi pola invarian di berbagai kondisi lingkungan. Hal ini meningkatkan konsep model dunia yang berevolusi sendiri dari sekadar beradaptasi dengan titik data baru menjadi memahami struktur kausal perubahan lingkungan.

Agen yang Berevolusi Sendiri

Meskipun model dunia memberikan agen pemahaman prediktif tentang lingkungannya, agen yang berevolusi sendiri adalah komponen inti yang memanfaatkan pemahaman ini untuk bertindak, belajar, dan meningkatkan diri. Evolusi agen terwujud adalah proses multifaset yang terjadi di sepanjang dua sumbu kritis dan saling terkait: penyempurnaan kecerdasan individunya melalui pengalaman fisik langsung, dan pengembangan kecerdasan kolektifnya melalui interaksi dengan agen lain. Sumbu pertama, Kognisi Berwujud yang Berevolusi Sendiri, mengeksplorasi bagaimana kecerdasan secara intrinsik terkait dengan bentuk fisik agen. Ini menyatakan bahwa evolusi agen didorong oleh lingkaran umpan balik dinamis antara tubuhnya, tindakannya, dan persepsinya tentang dunia. Sumbu kedua, Interaksi Dinamis yang Berevolusi Sendiri, memperluas konsep ini ke sistem multi-agen, di mana evolusi terjadi pada tingkat sosial. Di sini, agen harus belajar untuk berkoordinasi, berkomunikasi, dan mengembangkan konvensi bersama untuk mencapai tujuan kolektif. Bagian-bagian berikut menguraikan dua aspek mendasar dari evolusi agen ini.

Kognisi Wujud yang Berevolusi Sendiri

Teori kognisi wujud menantang pandangan klasik tentang kecerdasan yang terlepas dari tubuh, sebagai proses komputasional semata yang berada di dalam prosesor pusat. Sebaliknya, teori ini menyatakan bahwa kognisi pada dasarnya dibentuk oleh, dan muncul dari, interaksi dinamis dan timbal balik antara otak agen, tubuh fisiknya (morfologi dan kemampuan sensorimotor), dan lingkungannya. Dari perspektif ini, tubuh bukan hanya periferal pasif untuk merasakan dan bertindak, tetapi merupakan komponen aktif dan konstitutif dari proses kognitif itu sendiri. Prinsip bahwa "kognisi adalah untuk tindakan" menunjukkan bahwa kemampuan kognitif kita berevolusi untuk mendukung interaksi yang efektif dengan dunia,

dan dengan demikian, bentuk dan kemampuan fisik agen membatasi dan memungkinkan potensi kecerdasannya. Bagi AI wujud, ini berarti bahwa kecerdasannya harus didasarkan pada keberadaan fisik atau virtualnya. Evolusi kognisi wujud terjadi melalui setidaknya dua mekanisme yang saling terkait, yang memungkinkan agen untuk menyempurnakan interaksinya dengan dunia dari waktu ke waktu.

Peningkatan Diri melalui Persepsi Aktif. Belajar Cara Mempersepsi. Agen yang berwujud bukanlah pengamat pasif; ia secara aktif berpartisipasi dalam proses persepsi. Persepsi aktif mengacu pada kemampuan agen untuk secara cerdas mengontrol sensornya untuk memperoleh data yang lebih informatif dan menyelesaikan ambiguitas tentang keadaan dunia. Ini adalah bentuk evolusi diri di mana agen belajar untuk meningkatkan strategi akuisisi datanya sendiri. Misalnya, robot yang mencoba mengidentifikasi suatu objek mungkin belajar bahwa pandangan awalnya tidak informatif.

Melalui pembelajaran penguatan, ia dapat menemukan kebijakan tindakan, seperti menggerakkan kameranya atau memanipulasi objek, yang mengungkapkan fitur pembeda (misalnya, memutar cangkir untuk menemukan pegangannya). Interaksi lingkaran tertutup ini, di mana persepsi memandu tindakan dan tindakan memberi informasi pada persepsi, memungkinkan agen untuk mengembangkan sistem persepsi yang lebih efisien dan kuat yang disesuaikan dengan tugas dan lingkungannya.

Ko-evolusi Morfologi dan Kontrol. Bentuk evolusi diri yang lebih mendalam melibatkan perkembangan simultan dan berpasangan dari bentuk fisik agen (morfologi) dan kebijakan kontrolnya (kognisi/otak). Konsep ini merupakan tema sentral dalam robotika perkembangan dan evolusioner, yang mengambil inspirasi dari perkembangan ontogenetik organisme biologis. Di alam, tubuh dan otak suatu organisme berkembang bersama: perubahan morfologi menciptakan tantangan dan peluang baru bagi sistem kontrol, yang pada gilirannya mendorong pembelajaran dan adaptasi lebih lanjut. Misalnya, perubahan proporsi dan peningkatan kekuatan tubuh bayi manusia memberikan kurikulum alami untuk memperoleh keterampilan motorik seperti merangkak dan berjalan. Dalam sistem buatan, karya penting tentang evolusi makhluk virtual menunjukkan prinsip ini dengan menggunakan algoritma genetik untuk secara simultan mengembangkan rencana tubuh berbasis blok dan pengontrol jaringan saraf, yang mengarah pada munculnya strategi lokomosi yang beragam dan efektif yang tidak diprogram secara eksplisit tetapi ditemukan melalui evolusi simulasi. Secara lebih umum, algoritma ko-evolusi dapat mengoptimalkan morfologi robot dan parameter kontrol, memungkinkan penemuan hubungan sinergis di mana tubuh mengurangi kompleksitas komputasi yang signifikan dari pengendali. Dengan memperlakukan tubuh itu sendiri sebagai serangkaian parameter yang dapat dioptimalkan, agen dapat secara aktif membentuk lanskap evolusinya sendiri.

Riset dari kelompok kami memberikan wawasan penting ke dalam bidang ini. Makalah "AI Terwujud: Dari LLM ke Model Dunia" secara eksplisit membingkai interaksi yang terwujud sebagai penghubung penting antara pengetahuan abstrak, seperti yang ditemukan dalam LLM, dan kecerdasan fungsional yang mendasar. Lebih lanjut, kerangka kerja yang diusulkan dalam "Embodied-R: Kerangka Kerja Kolaboratif untuk Mengaktifkan Penalaran Spasial Terwujud

dalam Model Dasar melalui Pembelajaran Penguatan" adalah implementasi langsung dari kognisi yang berevolusi sendiri. Dalam karya ini, agen menggunakan pembelajaran penguatan untuk mempelajari cara berinteraksi dengan lingkungannya untuk mengaktifkan dan mendasarkan kemampuan penalaran spasialnya, secara efektif mengembangkan pemahaman kognitifnya sendiri tentang ruang melalui pengalaman yang terwujud. Terakhir, karya tentang "Meningkatkan Generalisasi Komposisional dalam Pembelajaran Lintas Perwujudan" membahas tantangan penting tentang bagaimana kemampuan kognitif yang dipelajari dapat digeneralisasikan di berbagai tubuh fisik, sebuah langkah kunci menuju penciptaan agen yang terwujud yang benar-benar adaptif dan serbaguna.

Interaksi Dinamis yang Berevolusi Sendiri

AI terwujud dan berevolusi sendiri juga meluas ke sistem multi-agen (MAS), di mana kecerdasan muncul dari interaksi yang kompleks dan dinamis di antara banyak entitas otonom. Dalam sistem seperti itu, agen harus belajar untuk mengoordinasikan tindakan mereka untuk mencapai tujuan bersama (kerja sama) atau untuk bersaing memperebutkan sumber daya yang terbatas (kompetisi) dalam lingkungan bersama. Tantangan utama terletak pada sifat lingkungan yang secara inheren tidak stasioner untuk setiap agen, karena kebijakan agen lain terus berubah dan beradaptasi. Hal ini memerlukan proses adaptasi timbal balik yang berkelanjutan dan evolusi strategi interaksi. Mekanisme utama yang mendorong evolusi interaksi dalam sistem multi-agen adalah komunikasi yang muncul, di mana agen belajar untuk mengembangkan sistem simbolik bersama mereka sendiri dari awal untuk memfasilitasi koordinasi. Dalam paradigma ini, agen tidak dibekali dengan protokol komunikasi yang telah diprogram sebelumnya. Sebaliknya, mereka belajar dalam kerangka kerja Pembelajaran Penguatan Multi-Agen (MARL), di mana saluran komunikasi merupakan bagian dari ruang aksi. Awalnya, agen dapat menghasilkan pesan acak, tetapi melalui penguatan pada imbalan tim bersama, sinyal yang mengarah pada hasil yang sukses diperkuat.

Kerangka kerja seperti Pembelajaran Antar-Agen yang Diperkuat (RIAL) dan Pembelajaran Antar-Agen yang Dapat Didiferensiasi (DIAL) menyediakan mekanisme untuk kemunculan ini. Sementara RIAL menggunakan pembelajaran Q-pembelajaran mendalam independen, DIAL meningkatkan hal ini dengan memanfaatkan pelatihan terpusat. Ia memperlakukan saluran komunikasi sebagai koneksi yang dapat didiferensiasi, memungkinkan gradien kesalahan untuk disebarkan kembali dari agen pendengar ke agen pembicara. Pembelajaran ujung-ke-ujung ini memberikan penugasan kredit kausal langsung yang memberi tahu pembicara bagaimana menyesuaikan pesannya untuk meningkatkan kinerja tim. Protokol yang dipelajari yang muncul dari komunikasi yang muncul pada dasarnya adalah konvensi sosial, yang merupakan pola perilaku yang sewenang-wenang tetapi stabil yang secara kolektif dibagi dan memfasilitasi koordinasi. Konvensi-konvensi ini membentuk dasar dari struktur sosial yang lebih kompleks dan memungkinkan agen untuk menavigasi dan menyelesaikan dilema sosial, yaitu situasi di mana tindakan yang rasional bagi individu mengarah pada hasil yang merugikan bagi kelompok. Karya fundamental tentang evolusi kerja sama menunjukkan bagaimana strategi timbal balik yang sederhana dapat mengarah pada kerja sama yang stabil dalam permainan berulang. Pembelajaran penguatan multi-agen

modern memperluas hal ini ke skenario yang lebih kompleks dan diperpanjang secara temporal yang dikenal sebagai dilema sosial sekuensial. Melalui penguatan pada imbalan bersama, perilaku komunikatif yang mengarah pada hasil yang saling menguntungkan diperkuat, sementara perilaku yang mengarah pada konflik ditekan.

Seiring waktu, pola perilaku komunikatif yang stabil ini menjadi norma sosial yang mengatur interaksi dalam masyarakat agen. Riset ekstensif dari kelompok kami tentang grafik dinamis dan jaringan saraf grafik menawarkan lensa formal dan kuat untuk menganalisis dan memodelkan interaksi multi-agen yang berkembang ini. Sistem multi-agen dapat secara alami direpresentasikan sebagai grafik dinamis, di mana agen adalah simpul dan tautan komunikasi atau pola interaksi mereka membentuk tepi. Keadaan sistem, termasuk keadaan internal agen dan protokol komunikasinya, berevolusi seiring waktu, menjadikannya masalah pembelajaran grafik dinamis. Makalah seperti "Penalaran Grafik Spasio-Temporal Dinamis untuk VideoQA" dan "Jaringan Perhatian Grafik Dinamis yang Terpisah untuk Rekomendasi Berurutan di Luar Distribusi" menyediakan arsitektur canggih untuk menangkap dan menalar tentang struktur relasional yang berevolusi tersebut. Lebih lanjut, tutorial tentang topik seperti "AI Generatif Multimoda dalam Lingkungan Dinamis dan Terbuka" dan pendalaman tentang pembelajaran penguatan multi-agen secara langsung membahas tantangan inti tentang bagaimana banyak agen dapat belajar untuk berkoordinasi dan beradaptasi di dunia yang kompleks dan terbuka, yang merupakan esensi dari evolusi interaksi dinamis.

Berdasarkan paradigma pendukung umum yang dibahas sebelumnya, penting juga untuk memeriksa kontribusi pengembangan secara spesifik yang telah memajukan komponen AI yang berevolusi sendiri. Di antaranya, karya Wang dan kawan kawan mewakili contoh berpengaruh yang menyediakan metode dan kerangka kerja konkret yang mendukung berbagai aspek evolusi diri. Tabel 4.2 merangkum kontribusi ini menurut tiga komponen yang diidentifikasi. Pandangan komparatif dalam Tabel 4.2 menyoroti bagaimana kontribusi individu dapat dipetakan ke komponen yang lebih luas dari AI yang berevolusi sendiri. Dengan menyelaraskan kemajuan metodologis dengan pemodelan dunia, kognisi yang terwujud, dan interaksi dinamis, karya-karya ini menggambarkan jalur konkret di mana paradigma teoretis dapat diwujudkan dalam pengembangan praktis.

Tabel 4.2 Komponen AI terwujud berevolusi mandiri menyoroti peningkatan pemodelan dunia, kognisi terwujud, dan interaksi dinamis

Komponen	Publikasi/tutorial terkait	Kontribusi utama untuk evolusi diri
Model Dunia yang Berkembang Sendiri	Jaringan Saraf Graf Dinamis di Bawah Pergeseran Distribusi Spasio-Temporal	Menyediakan kerangka kerja untuk memperbarui model dunia dengan memperlakukan dinamika lingkungan sebagai grafik di bawah pergeseran distribusi, memungkinkan adaptasi berkelanjutan
Kognisi Terwujud yang Berkembang Sendiri	Embodied-R: Kerangka Kerja Kolaboratif untuk Mengaktifkan Penalaran Spasial yang Terwujud dalam	Mendemonstrasikan bagaimana agen dapat mengembangkan kemampuan kognitifnya (penalaran spasial) melalui interaksi aktif dan terwujud dengan

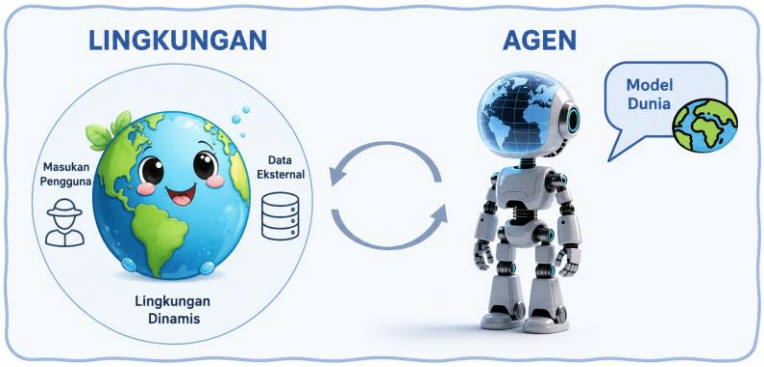
	Model Dasar melalui Pembelajaran Penguatan	lingkungan, yang didorong oleh pembelajaran penguatan
Interaksi Dinamis yang Berkembang dengan Sendirinya	Jaringan Modular Spasio-Temporal Dinamis untuk Tanya Jawab Video	Memperkenalkan jaringan modular yang secara dinamis melakukan penalaran atas interaksi spasial-temporal, menyediakan mekanisme untuk memodelkan hubungan yang berkembang antara agen atau objek dalam suatu adegan

BAB 5
MODEL DUNIA YANG BEREVOLUSI SENDIRI

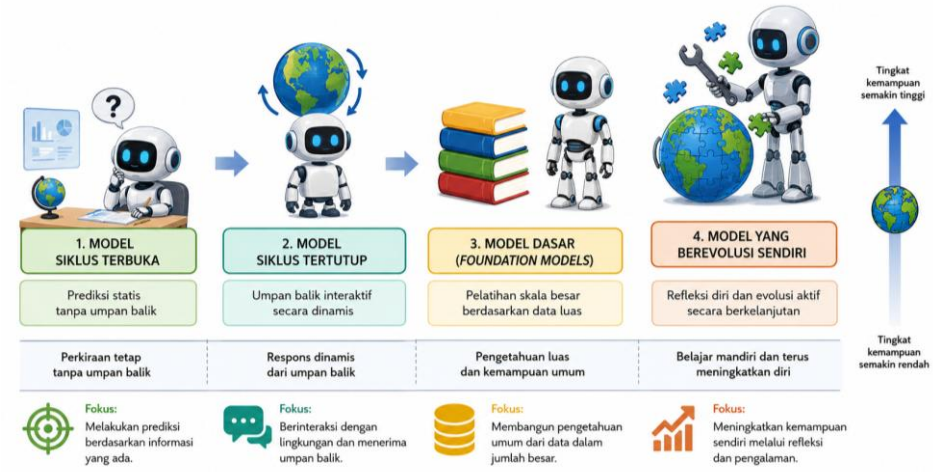
5.1 PENDAHULUAN

Model dunia yang berevolusi sendiri (SEWM) merupakan kemajuan mendasar dalam kecerdasan buatan, yang memungkinkan agen untuk secara otonom memperbaiki representasi internal dan kemampuan prediktif mereka melalui interaksi berkelanjutan dengan lingkungan. Bab ini menyediakan kerangka kerja komprehensif dengan memformalkan SEWM menjadi empat komponen inti: Model (meliputi persepsi, tugas, kebijakan, aktor, dan imbalan), Memori, Alat, dan Pengoptimal Dinamika. Secara matematis, model dunia yang berevolusi sendiri bertujuan mempelajari fungsi dinamika $P(s_{t+1} \mid s_t, a_t, M, D, L, P)$ yang mencapai kesetaraan fungsional dengan lingkungan sebenarnya. Mekanisme evolusi diri digerakkan oleh optimasi putaran tertutup melalui Pengoptimal Dinamika. Bab ini secara sistematis membahas evolusi diri di dalam setiap komponen, termasuk evolusi persepsi, tugas, kebijakan, aktor, dan imbalan pada model, serta evolusi memori, alat, dan Pengoptimal Dinamika. Bab ini juga membahas evolusi diri antar-komponen, seperti evolusi diri bersama model-memori, model-alat, dan model-Pengoptimal Dinamika. Formalisme terstruktur ini mendukung pengembangan model dunia yang adaptif, tangguh, dan mampu melakukan generalisasi dalam lingkungan dinamis.

Model Dunia adalah representasi internal agen dan prediksi masa depan tentang dunia nyata, seperti yang ditunjukkan pada Gambar 5.1, yang dapat mengabstraksikan model dinamis lingkungan objektif dan memahami prinsip-prinsip yang mendasari dunia nyata. Dengan mengompresi masukan sensorik berdimensi tinggi ke dalam ruang laten terstruktur, model dunia memfasilitasi penalaran yang efisien, perencanaan jangka panjang, dan imajinasi kontrafaktual. Kemampuan ini sangat penting dalam aplikasi dunia nyata seperti mengemudi otonom, manipulasi robot, dan sistem AI interaktif, di mana agen harus beroperasi di bawah ketidakpastian dan pengamatan parsial. Paradigma pelatihan untuk model dunia telah mengalami evolusi yang signifikan, seperti yang ditunjukkan pada Gambar 5.2. Model dunia putaran terbuka awal dilatih pada himpunan data tetap dan tidak memiliki kemampuan untuk berinteraksi atau beradaptasi dengan lingkungannya. Munculnya model dunia putaran tertutup menandai pergeseran menuju pembelajaran daring, di mana agen dapat memperbarui model mereka berdasarkan umpan balik waktu nyata. Baru-baru ini, model dunia dasar telah muncul, memanfaatkan pelatihan awal skala besar pada data multimoda yang beragam untuk memperoleh representasi yang luas dan dapat digeneralisasikan. Saat ini, kita berada di garis depan model dunia yang berevolusi sendiri, yang dapat secara otonom memperbaiki struktur internalnya, belajar dari pengalamannya sendiri, dan terus meningkat tanpa pengawasan manusia.



Gambar 5.1 Konsep model dunia



Gambar 5.2 Evolusi paradigma pelatihan model dunia

Model dunia yang berevolusi sendiri didefinisikan tidak hanya oleh kemampuannya untuk memprediksi dan mensimulasikan, tetapi juga oleh kapasitasnya untuk peningkatan diri secara otonom. Mengambil inspirasi dari ilmu kognitif dan ilmu saraf, model tersebut mewujudkan prinsip-prinsip seperti neuroplastisitas, pengkodean prediktif, dan metakognisi. Model ini terlibat dalam siklus iteratif akuisisi pengalaman, penyempurnaan, dan integrasi, mirip dengan proses kognitif manusia yang dijelaskan oleh teori perkembangan kognitif Piaget. Mekanisme kuncinya meliputi plastisitas struktural, pertumbuhan modular, refleksi diri, konsolidasi pengetahuan, dan kelupaan adaptif, yang semuanya memungkinkan model untuk tetap efisien dan relevan dalam lingkungan yang berubah.

Secara formal, model dunia yang berevolusi sendiri dapat dicirikan sebagai sistem dinamis yang mengoptimalkan arsitektur dan parameternya sendiri melalui interaksi berkelanjutan dengan lingkungan. Model ini mengintegrasikan teknik dari pembelajaran penguatan, pembelajaran berkelanjutan, meta pembelajaran, pembelajaran kurikulum, dan pembelajaran mesin otonom untuk mencapai pertumbuhan dan adaptasi yang berkelanjutan. Tujuannya adalah untuk menciptakan agen yang tidak hanya berkinerja baik pada tugas-tugas yang telah ditentukan sebelumnya, tetapi juga menemukan tujuan baru, menyempurnakan strategi mereka, dan mengembangkan kemampuan mereka dari waktu ke waktu.

Bab ini bertujuan untuk menyediakan kerangka kerja komprehensif untuk memahami dan mengembangkan model dunia yang berevolusi sendiri. Pertama, model dunia saat ini hanya memiliki definisi fungsional: Apa pun yang memenuhi persyaratan "memahami dunia dan memprediksi masa depan" dianggap sebagai model dunia. Bab ini adalah yang pertama membangun definisi model dunia yang terstruktur dan diformalkan serta mengidentifikasi komponen-komponen kuncinya, seperti model, memori, alat, dan pengoptimal dinamika, yang dapat sangat mendorong pengembangan dan iterasi sistematis model dunia. Kedua, bab ini mengatur teori dan metode untuk evolusi diri dalam komponen-komponen, seperti evolusi diri model, evolusi diri memori, evolusi diri alat, dan evolusi diri pengoptimal dinamika. Secara khusus, untuk evolusi diri model, kita mempertimbangkan bagaimana seluruh proses model berevolusi, termasuk bagaimana persepsi lingkungan, dekomposisi tugas, pengambilan keputusan kebijakan, pemilihan tindakan, dan optimasi fungsi imbalan berevolusi sendiri. Terakhir, bab ini mengorganisir teori dan metode untuk evolusi diri antar komponen, seperti evolusi diri model-memori bersama, evolusi diri model-alat bersama, evolusi diri model-pengoptimal dinamika bersama, dan lainnya. Dengan menjembatani kesenjangan antara fondasi teoretis dan implementasi praktis, kami berharap dapat mempercepat pengembangan sistem AI yang benar-benar otonom, adaptif, dan dapat digeneralisasi yang mampu berkembang di dunia yang terus berubah.

5.2 LANDASAN MODEL DUNIA YANG BEREVOLUSI SENDIRI

Konsep model dunia, yang sering digambarkan secara informal sebagai representasi internal agen terhadap dunia objektif dan mekanisme prediksinya di masa depan, telah menjadi pusat kecerdasan buatan dan ilmu kognitif. Namun, definisinya sebagian besar tetap fungsional, tanpa dekomposisi terstruktur dan terformal yang memfasilitasi rekayasa sistematis dan peningkatan iteratif. Bagian ini mengatasi kesenjangan ini dengan mengusulkan definisi baru dan komprehensif tentang Model Dunia yang Berevolusi Sendiri. Kami mendekonstruksi model dunia menjadi empat modul inti yang saling berinteraksi: Model, Memori, Alat, dan Pengoptimal Dinamika. Yang terpenting, kami memperkenalkan konsep evolusi diri sebagai properti intrinsik, di mana model terus menerus menyempurnakan fungsi dinamika internalnya melalui pengalaman. Tujuan utama kerangka kerja ini adalah untuk mempelajari Fungsi Dinamika yang secara probabilistik memetakan keadaan dan tindakan saat ini ke keadaan berikutnya, yang dikondisikan pada modul internal. Kami memformalkan tujuan tersebut sebagai pencapaian kesetaraan fungsional, di mana dinamika yang dipelajari konsisten dengan dinamika lingkungan sebenarnya untuk semua kemungkinan tindakan. Formalisasi ini menyediakan cetak biru dasar untuk pengembangan sistem cerdas yang lebih tangguh, dapat digeneralisasi, dan otonom.

Mengapa Diperlukan Formalisasi

Gagasan model internal berakar kuat di berbagai bidang, mulai dari teori kontrol, dengan konsep seperti filter Kalman, hingga psikologi, melalui peta kognitif Tolman. Dalam pembelajaran penguatan, perbedaan antara metode bebas model dan berbasis model bergantung pada ada atau tidaknya model dunia. Karya-karya perintis seperti arsitektur Dyna

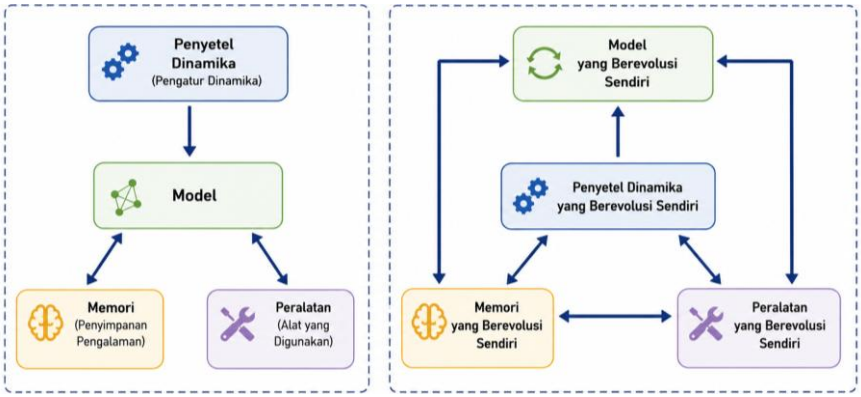
Sutton mengintegrasikan model dunia untuk perencanaan. Baru-baru ini, pendekatan pembelajaran mendalam telah menghasilkan model dunia yang mengesankan, seperti yang digunakan dalam seri agen Dreamer, yang mempelajari representasi lingkungan yang ringkas dari piksel dan menggunakannya untuk perencanaan di ruang laten. Terlepas dari kemajuan ini, definisi formal yang konsisten dan terperinci dari model dunia itu sendiri seringkali tidak ada. Model tersebut biasanya implisit dalam arsitektur, sebagai produk sampingan dari pelatihan jaringan saraf pada tugas prediksi. Kurangnya formalisasi eksplisit ini menimbulkan tantangan. Hal ini menyulitkan untuk membandingkan arsitektur model dunia yang berbeda, untuk mendiagnosis kegagalan, untuk secara sistematis menggabungkan kemampuan baru seperti penggunaan alat atau memori episodik, dan untuk merekayasa komponennya secara independen. Karya kami mengambil inspirasi dari pendekatan-pendekatan sebelumnya, tetapi berupaya untuk mengabstraksikan elemen-elemen umum mereka ke dalam kerangka kerja yang umum, modular, dan terdefinisi secara formal. Dengan secara eksplisit menyebutkan dan mendefinisikan komponen-komponen seperti Memori, Alat, dan Pengoptimal Dinamika, kami menyediakan kosakata umum dan cetak biru struktural untuk pengembangan di masa mendatang.

Komponen Kunci

Dalam upaya mencapai Kecerdasan Buatan Umum (AGI), gagasan tentang model dunia menjadi landasan. Model ini merangkum kemampuan agen untuk tidak hanya bereaksi terhadap rangsangan, tetapi juga memiliki simulasi internal lingkungannya. Model internal ini memungkinkan pemahaman, penalaran, dan prediksi, yang sangat penting untuk perencanaan dan perilaku yang canggih. Definisi tradisional, meskipun secara konseptual masuk akal, seringkali kurang jelas. Deskripsi seperti "abstraksi dinamika lingkungan" atau "sistem yang memahami aturan dunia" memberikan intuisi tingkat tinggi tetapi menawarkan sedikit panduan untuk implementasi atau analisis yang ketat. Mereka memperlakukan model dunia sebagai kotak hitam monolitik yang memenuhi persyaratan fungsional, yaitu pemahaman dan prediksi.

Subbagian ini menyajikan penyimpangan signifikan dari perspektif ini. Kami berpendapat bahwa agar model dunia benar-benar efektif dan terukur, terutama dalam lingkungan yang kompleks dan terbuka, model tersebut harus diuraikan secara arsitektur dan dilengkapi dengan kemampuan untuk meningkatkan diri. Oleh karena itu, kami mengusulkan definisi formal dari Model Dunia yang Berevolusi Sendiri. Definisi ini melampaui deskripsi fungsional menuju deskripsi struktural dan operasional. Kontribusi utama adalah identifikasi dan spesifikasi rinci dari empat komponen fundamental yang membentuk model dunia, dan formalisasi interaksi mereka yang bertujuan untuk mempelajari fungsi dinamika yang tepat. Sifat evolusi diri muncul dari interaksi berkelanjutan antara komponen-komponen ini, khususnya melalui peran Pengoptimal Dinamika. Kerangka kerja ini dimaksudkan untuk berfungsi sebagai fondasi terstruktur, yang mendorong pengembangan sistematis, pembandingan kinerja, dan analisis teoretis model dunia. Model Dunia yang Berevolusi Sendiri yang diusulkan didefinisikan sebagai sistem terintegrasi yang terdiri dari empat modul inti.

Interaksi antara modul-modul ini diilustrasikan pada Gambar 5.3 dan membentuk dasar untuk operasi dan evolusi model.



Gambar 5.3 Komponen-komponen model dunia yang berevolusi sendiri

Inti Model $(\mathbf{M})$

Inti Model merupakan mesin prediksi utama model dunia yang dilambangkan sebagai M. Inti Model bertanggung jawab mempertahankan representasi keadaan saat ini dan menjalankan fungsi dinamika. Inti Model dapat diuraikan lebih lanjut menjadi beberapa submodul khusus:

- Persepsi (O): Submodul ini memproses masukan sensorik mentah dari lingkungan dan mengubahnya menjadi representasi keadaan internal s_t. Submodul ini menangani tugas seperti pengenalan objek, estimasi keadaan, dan ekstraksi fitur.
- Tugas (T): Modul ini mengodekan tujuan atau sasaran agen saat ini. Modul ini memodulasi perilaku Kebijakan dan memberikan konteks tentang aspek dinamika dunia yang paling relevan.
- Kebijakan (π): Kebijakan merupakan fungsi yang memetakan keadaan saat ini s_t dan tugas T ke distribusi atas tindakan a_t. Kebijakan ini merepresentasikan strategi agen saat ini untuk mencaptugas AInya.
- Aktor (A): Aktor merupakan komponen yang mengambil sampel tindakan tertentu a_t dari distribusi kebijakan dan mengeksekusinya di lingkungan atau dalam simulasi internal.
- Imbalan (R): Modul ini memprediksi imbalan yang diharapkan r_{t+1} sebagai hasil dari pengambilan tindakan a_t pada keadaan s_t. Modul ini merupakan komponen penting untuk mengevaluasi konsekuensi tindakan.

Integritas Inti Model bergantung pada interaksi yang mulus dari submodul-submodul ini. Persepsi menginformasikan keadaan, Tugas memandu Kebijakan, Aktor mengeksekusi tindakan, dan modul Imbalan mengevaluasi hasilnya.

Memori $(\mathbf{D})$

Modul Memori, yang dilambangkan sebagai D, berfungsi sebagai repositori pengalaman masa lalu agen. Modul ini bukan penyimpanan tunggal, tetapi kemungkinan terdiri atas beberapa subsistem:

- Memori Episodik**:** Merekam peristiwa masa lalu tertentu dalam bentuk $(s_t, a_t, r_{t+1}, s_{t+1})$.
- Memori Semantik: Menyimpan pengetahuan umum dan fakta tentang dunia, yang diabstraksikan dari pengalaman spesifik.
- Memori Prosedural: Mengkodekan keterampilan dan rutinitas untuk menggunakan alat dan melakukan tindakan.

Modul Memori menyediakan fondasi data tempat model dunia dibangun dan disempurnakan. Modul ini memungkinkan agen belajar dari sejarah, menghindari kesalahan masa lalu, dan mengingat strategi yang berhasil. Pengoptimal Dinamika sangat bergantung pada D untuk melatih Model Inti.

Alat $(\mathbf{L})$

Modul Alat, yang dilambangkan sebagai L, mewakili serangkaian kemampuan eksternal atau internal yang dapat dipanggil agen untuk berinteraksi dengan dunia atau membantu proses penalarannya sendiri. Alat memperluas kemampuan asli agen. Contohnya meliputi:

- Alat fisik (misalnya, lengan robot yang menggunakan kunci inggris).
- Alat komputasi (misalnya, kalkulator, sistem kueri basis data).
- Alat konseptual (misalnya, simulator fisika, model bahasa).

Aspek kunci dari alat adalah bahwa pengaruhnya terhadap dunia harus dapat dipelajari dan diintegrasikan ke dalam fungsi dinamika secara keseluruhan. Submodul Kebijakan harus belajar untuk memutuskan tidak hanya tindakan primitif tetapi juga kapan dan bagaimana menggunakan alat yang tersedia.

Pengoptimal Dinamika $(\mathbf{P})$

Pengoptimal Dinamika adalah mesin evolusi diri yang dilambangkan sebagai P. Komponen metakognitif ini bertanggung jawab meningkatkan akurasi prediksi Model Inti. Fungsi utamanya adalah membandingkan secara terus-menerus prediksi Model Inti dengan hasil aktual yang diamati dari lingkungan atau diambil dari memori. Selanjutnya, komponen ini menghitung sinyal kerugian berdasarkan kesalahan prediksi, misalnya perbedaan antara keadaan berikutnya yang diprediksi, s_{t+1}, dan keadaan aktual, atau perbedaan antara imbalan yang diprediksi dan imbalan aktual. Sinyal tersebut digunakan oleh algoritma optimasi untuk memperbarui parameter submodul Model Inti. Proses penyempurnaan iteratif ini memungkinkan model dunia menjadi semakin akurat dan komprehensif dari waktu ke waktu.

Formalisasi Matematis

Tujuan utama dari Model Dunia yang Berevolusi Sendiri adalah untuk membangun Fungsi Dinamika yang sangat akurat. Fungsi ini adalah inti matematis dari Model Inti. Kami mendefinisikannya sebagai distribusi probabilitas bersyarat:

$$P(s_{t+1} \mid s_t, a_t, M, D, L, P)$$

Fungsi ini memprediksi distribusi probabilitas keadaan berikutnya s_{t+1} berdasarkan keadaan saat ini s_t, tindakan yang diambil a_t, serta kondisi Model M saat ini, isi Memori D, Alat L yang

tersedia, dan Pengoptimal Dinamika P. Formulasi ini mengakui bahwa dinamika dunia tidak absolut tetapi dipersepsikan dan diinterpretasikan melalui aparatus kognitif agen saat ini.

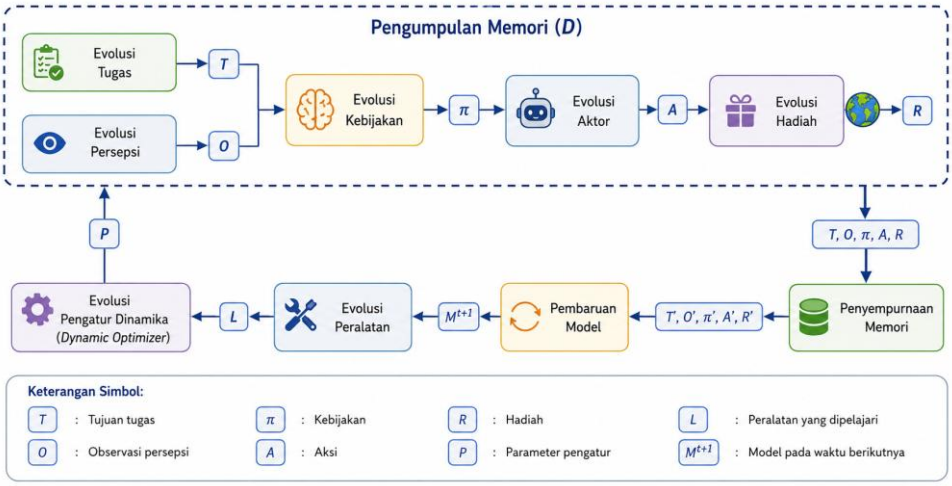
Tujuan pembelajaran adalah memastikan bahwa fungsi dinamika internal selaras secara sempurna dengan dinamika lingkungan sebenarnya, $P_{\text{env}}(s_{t+1} \mid s_t, a_t)$. Kriteria formal keberhasilannya adalah kesetaraan fungsional. Model yang dipelajari dinyatakan setara secara fungsional dengan lingkungan sebenarnya jika, untuk setiap keadaan s yang dapat direpresentasikan oleh agen dan setiap tindakan a dalam domain tindakan yang mungkin, kondisi berikut terpenuhi:

$$\forall a \in \mathcal{A}, \quad P(s_{t+1} \mid s_t, a) = P_{\text{env}}(s'_{t+1} \mid s_t, a)$$

Dengan kata lain, untuk setiap tindakan, distribusi prediksi atas keadaan selanjutnya (dan akibatnya, imbalan) yang dihasilkan oleh model internal identik dengan distribusi yang dihasilkan oleh lingkungan sebenarnya. Mencapai kesetaraan ini berarti simulasi internal agen terhadap dunia sama akuratnya dengan dunia itu sendiri, memungkinkan prediksi sempurna dan perencanaan optimal.

Mekanisme Evolusi Diri

Istilah evolusi diri merupakan unsur utama dalam definisi ini. Evolusi digerakkan oleh operasi Pengoptimal Dinamika secara terus-menerus dalam putaran tertutup dengan modul lain. Seperti ditunjukkan pada Gambar 5.4, (1) agen berinteraksi dengan lingkungan menggunakan Model Inti saat ini; (2) tuple pengalaman baru $(s_t, a_t, r_{t+1}, s_{t+1})$ disimpan dalam Memori D; (3) Pengoptimal Dinamika mengambil sampel sekumpulan pengalaman dari D; (4) komponen tersebut menghitung kesalahan prediksi dengan membandingkan prediksi keadaan dan imbalan dari Model Inti dengan hasil aktual; (5) sinyal kesalahan digunakan untuk memperbarui parameter jaringan Persepsi, Imbalan, dan Kebijakan di dalam Model Inti melalui propagasi balik atau teknik optimasi lainnya; dan (6) Model Inti yang telah diperbarui digunakan pada siklus interaksi berikutnya.



Gambar 5.4 Mekanisme evolusi diri model dunia

5.3 EVOLUSI DIRI DI DALAM KOMPONEN

Subbagian ini mengorganisir teori dan metode untuk evolusi diri dalam komponen, seperti evolusi diri model, evolusi diri memori, evolusi diri alat, dan evolusi diri pengoptimal dinamika. Secara khusus, untuk evolusi diri model, bagian ini mempertimbangkan bagaimana seluruh proses model berevolusi.

Evolusi Diri Model

Bagian ini mempertimbangkan bagaimana seluruh proses model berevolusi, termasuk bagaimana persepsi lingkungan, dekomposisi tugas, pengambilan keputusan kebijakan, pemilihan tindakan, dan optimasi fungsi imbalan berevolusi sendiri.

Evolusi Persepsi

Dalam model dunia konvensional, modul persepsi biasanya diimplementasikan sebagai pengode statis, seperti jaringan konvolusional atau pengode otomatis variasional, yang mengompres pengamatan mentah menjadi keadaan laten berdimensi rendah. Desain ini mengasumsikan bahwa vektor laten yang kompak cukup untuk mendukung prediksi dan kontrol dinamika hilir. Meskipun efektif dalam domain sederhana, modul persepsi yang tidak berevolusi tersebut terbatas kemampuannya untuk melakukan generalisasi di berbagtugas AI dan lingkungan. Mereka terutama mengoptimalkan untuk rekonstruksi atau prediksi satu langkah, seringkali menggabungkan beberapa faktor variasi ke dalam representasi monolitik. Akibatnya, mereka kesulitan untuk menangkap struktur tingkat objek, untuk mengintegrasikan berbagai modalitas, atau untuk mempertahankan memori jangka panjang yang koheren, sehingga membatasi imajinasi dan kemampuan perencanaan.

Oleh karena itu, riset terbaru telah beralih ke arah evolusi modul persepsi itu sendiri. Alih-alih tetap menjadi kompresor fitur statis, persepsi dalam model dunia sekarang diperlakukan sebagai komponen dinamis yang menyesuaikan representasinya saat agen menghadapi keadaan, tugas, dan lingkungan baru. Tujuan prediktif yang diawasi sendiri mengembangkan persepsi menjadi ruang laten yang berorientasi masa depan; metode berpusat pada objek mengembangkan persepsi menjadi representasi yang terstruktur dan dapat diinterpretasikan; pemodelan temporal memperluas persepsi menjadi dinamika jangka panjang; dan desain multimoda yang sadar geometri mengembangkan persepsi menjadi ruang laten lintas sensor yang konsisten secara spasial. Subbagian berikut menguraikan keempat lintasan ini dan karya-karya representatifnya.

Prediktif yang Diawasi Sendiri. Model dunia awal menunjukkan bahwa persepsi dapat berevolusi dengan mempelajari ruang laten prediktif. Model Dunia asli dan PlaNet memperkenalkan dinamika laten stokastik-deterministik kompak yang dilatih dengan tujuan variasional, memungkinkan peluncuran multi-langkah sepenuhnya di ruang laten. Ini menetapkan prinsip bahwa persepsi harus melayani imajinasi dengan mengkodekan keadaan yang kompak namun prediktif. Seri Dreamer memajukan pendekatan ini dengan laten diskrit, peningkatan regularisasi, dan tujuan pelatihan yang seimbang, secara signifikan meningkatkan stabilitas dan efisiensi sampel. Metode berbasis eksplorasi seperti Plan2Explore memperluas persepsi prediktif dengan memaksimalkan ketidakpastian epistemik dan perolehan informasi, memastikan bahwa representasi laten secara adaptif mencakup keadaan yang tidak terlihat.

Terstruktur Berbasis Objek. Arah riset lain mengembangkan persepsi menjadi representasi berbasis objek dan relasi yang terstruktur. Alih-alih vektor fitur yang saling terkait, adegan diuraikan menjadi entitas dan interaksinya. SAVi/SAVi++ memperluas perhatian slot ke prediksi video, memungkinkan pelacakan objek temporal. Model Dunia Terstruktur Slot (SSWM) menggabungkan perhatian slot dengan dinamika grafik untuk memodelkan interaksi objek secara eksplisit, sementara SOLD mempelajari dinamika laten berpusat pada objek tanpa pengawasan, mendukung generalisasi kombinatorial ke konfigurasi objek yang tidak terlihat. Kemajuan terbaru seperti Dyn-O memperkenalkan slot dinamis yang berevolusi seiring waktu, dan Jeong dan kawan kawan menyelaraskan persepsi dengan kondisi bahasa alami, menjembatani persepsi terstruktur dengan semantik yang relevan dengan tugas. Bersama-sama, karya-karya ini menyoroti evolusi persepsi dari pengkodean tingkat piksel ke struktur laten tingkat objek yang dapat diinterpretasikan.

Temporal Urutan Panjang. Trajektori ketiga membahas tantangan pemodelan ketergantungan rentang panjang dalam persepsi. Model laten rekuren seperti RSSM seringkali mengalami keterbatasan kapasitas memori. Untuk mengatasi hal ini, IRIS menggabungkan pengodean otomatis diskrit dengan Transformator, menghasilkan model dunia yang efisien sampel dengan memori yang diperluas. STORM mengintegrasikan variabel laten stokastik ke dalam tulang punggung Transformator, menyeimbangkan ekspresivitas dan stabilitas untuk pemodelan urutan panjang. TWISTER lebih lanjut meningkatkan pemodelan rentang dengan pengkodean prediktif kontrastif yang dikondisikan aksi, sementara Model Dunia Video SSM Berkonteks Panjang memperluas cakupan temporal menggunakan arsitektur ruang keadaan dengan inferensi waktu linier. Kemajuan ini menunjukkan bagaimana persepsi berevolusi menjadi modul temporal yang mampu mempertahankan imajinasi yang koheren selama ratusan langkah.

Multimoda dan Sadar Geometri. Akhirnya, persepsi berevolusi dengan mengintegrasikan sinyal multimoda dan geometri spasial. Aplikasi dunia nyata seperti mengemudi otonom membutuhkan penggabungan isyarat visual, LiDAR, dan geometris. UniWorld secara bersamaan mengkodekan visi dan LiDAR ke dalam ruang laten terpadu untuk memprediksi okupansi 4D, yang tahan terhadap oklusi dan derau sensor. OccWorld dan Drive-OccWorld lebih lanjut menyempurnakan pendekatan ini dengan secara eksplisit memprediksi aliran okupansi dan menyelaraskan isyarat semantik dan gerak. UniFuture memperkenalkan ruang laten ganda untuk citra dan kedalaman dengan interaksi lintas modal multi-skala, menyatukan persepsi dan prediksi untuk peramalan lintasan dan penyelesaian semantik. DIO menguraikan aliran okupansi menjadi struktur laten tingkat instans, meningkatkan komposisi dan interpretasi. Karya-karya ini secara kolektif menggambarkan bagaimana persepsi berevolusi dari pengode modalitas tunggal menjadi ruang laten spasial-temporal yang sadar geometri, memungkinkan perencanaan yang kuat di lingkungan yang kompleks.

Evolusi Tugas

Dalam model dunia konvensional, tugas biasanya diperlakukan sebagai sesuatu yang tetap dan didefinisikan secara eksternal. Lingkungan menentukan tujuan, seperti mencapai lokasi target atau memaksimalkan imbalan skalar, dan model dunia hanya dilatih untuk

memprediksi transisi keadaan atau imbalan di bawah tujuan yang telah ditentukan sebelumnya ini. Dalam pengaturan ini, tugas adalah kendala eksternal daripada komponen intrinsik dari model. Formulasi ini memiliki beberapa keterbatasan. Tugas tetap membatasi kemampuan model untuk melakukan generalisasi di berbagai tujuan, mencegah penemuan keterampilan baru, dan tidak menawarkan mekanisme adaptasi ketika tugas baru ditemui. Akibatnya, model dunia yang dilatih di bawah definisi tugas statis seringkali kurang fleksibel dan kesulitan dalam skenario terbuka atau multi-tugas.

Kemajuan terbaru telah mendefinisikan ulang tugas sebagai entitas dinamis yang berevolusi bersama dengan komponen lain dari model dunia. Kami menyebut arah ini sebagai evolusi diri tugas. Tidak seperti tolok ukur statis, tugas yang berevolusi sendiri tidak ditentukan sebelumnya, melainkan muncul melalui pengkondisian pada variabel kontekstual, melalui pembangkitan tujuan baru secara otomatis, atau melalui mekanisme adaptif yang memungkinkan satu model untuk menyesuaikan diri di berbagai tujuan. Pendekatan ini dapat dikelompokkan menjadi tiga kategori utama. Pengkondisian tugas menyuntikkan informasi tugas langsung ke dalam ruang laten model dunia, memungkinkan imajinasi dan perencanaan untuk beradaptasi dengan beragam tujuan. Pembangkitan tugas memperluas distribusi tugas dengan secara otomatis membangun tantangan atau kurikulum baru, memastikan bahwa kemampuan agen berkembang seiring dengan kompleksitas tugas. Terakhir, adaptasi tugas berfokus pada memungkinkan satu model dunia untuk melakukan generalisasi di berbagtugas AI dan untuk mentransfer secara efektif ke tujuan baru. Subbagian berikut meninjau setiap kategori secara detail dan menyoroti metode representatif.

Pengkondisian Tugas. Kategori pertama mengkodekan informasi tugas langsung dalam model dunia, memungkinkan satu arsitektur untuk secara fleksibel menangani beragam tujuan. Alih-alih melatih model terpisah per tugas, variabel kontekstual seperti tujuan, instruksi bahasa, atau prioritas fisik disematkan ke dalam proses imajinasi laten. LEXA merepresentasikan tugas sebagai tujuan laten dengan mempelajari jarak temporal dalam imajinasi, memberikan sinyal imbalan yang mencerminkan kesulitan kontrol. Pengkondisian berbasis bahasa telah memperluas paradigma ini lebih lanjut. LUMOS mengintegrasikan perintah bahasa alami ke dalam imajinasi laten untuk mendukung manipulasi robotik jangka panjang, sementara LIMT memperoleh penyematan semantik dari model bahasa yang telah dilatih sebelumnya untuk mengarahkan prediksi dan kontrol multi-tugas. Pendekatan lain memperluas pengkondisian ke fisika dan pengetahuan yang telah dilatih sebelumnya. Model dunia yang diinformasikan fisika seperti PIN-WM dan RoboScape membatasi keadaan laten untuk menghormati konsistensi fisik, memungkinkan transfer yang lebih baik ke pengaturan dunia nyata. Model yang selaras dengan pelatihan sebelumnya seperti LaDi-WM menggabungkan prioritas semantik dari model dasar visi, memperkaya representasi tugas dengan pengetahuan tingkat tinggi. Karya-karya ini menunjukkan bahwa pengkondisian mengubah tugas dari kendala eksternal menjadi sinyal intrinsik yang membentuk pembelajaran model dunia.

Generasi Tugas. Kategori kedua berfokus pada evolusi ruang tugas itu sendiri daripada pengkondisian pada tujuan yang ada. Dengan terus menerus menghasilkan tantangan baru,

pendekatan ini memastikan bahwa model dan kebijakan dunia berkembang seiring dengan kompleksitas tugas. Platform terprogram seperti XLand mendefinisikan distribusi besar tugas yang dihasilkan secara prosedural, menunjukkan bahwa keragaman sangat penting untuk akuisisi keterampilan yang terbuka. Metode yang lebih baru membangun kurikulum tugas secara dinamis. Hill dan kawan kawan membangun tugas baru dengan memberikan tantangan kepada agen, sehingga menghasilkan kurikulum otomatis yang skalanya meningkat seiring dengan peningkatan agen. Baik dalam pendekatan prosedural maupun adversarial, evolusi tugas muncul melalui perluasan distribusi tugas itu sendiri, memungkinkan model dunia untuk mengembangkan kompetensi yang luas di berbagai domain.

Adaptasi Tugas. Kategori ketiga menekankan bagaimana satu model dunia dapat melakukan generalisasi di berbagtugas AI dan dengan cepat beradaptasi dengan tujuan baru. Alih-alih membuat atau mengkondisikan tugas, metode ini menanamkan kesadaran tugas ke dalam proses pelatihan. TEMPO memperkenalkan skema pembobotan meta yang memprioritaskan data yang lebih informatif untuk tujuan saat ini, meningkatkan efisiensi dalam pembelajaran multi-tugas. WAKER mengadopsi desain kurikulum tanpa imbalan, memilih contoh lingkungan yang meminimalkan kesalahan prediksi, sehingga menghasilkan model dunia yang kuat yang dapat ditransfer secara efektif ke tugas-tugas selanjutnya. Kemajuan empiris lebih lanjut mendukung perspektif ini: Model Dunia Multi-tugas yang Terpisah menunjukkan bahwa pelatihan di berbagtugas AI mengarah pada ruang laten yang terpisah, meningkatkan generalisasi lintas tugas. Dengan demikian, adaptasi tugas menyoroti bagaimana tugas itu sendiri dapat bertindak sebagai bias induktif yang membentuk pelatihan model dunia, mendukung transfer yang kuat dan perencanaan yang fleksibel di berbagai domain.

Evolusi Kebijakan

Dalam model dunia standar, modul kebijakan sering diwujudkan sebagai jaringan saraf statis, seperti perceptron multi-layer atau kepala berulang, yang memetakan keadaan laten langsung ke distribusi tindakan. Kebijakan ini biasanya dioptimalkan menggunakan metode pembelajaran penguatan, misalnya gradien kebijakan atau algoritma aktor-kritik. Meskipun berfungsi di lingkungan sederhana, kebijakan statis tersebut memiliki beberapa keterbatasan. Kebijakan tersebut sangat bergantung pada imbalan yang jarang atau tertunda, yang menyebabkan efisiensi sampel yang buruk, dan hanya memanfaatkan kapasitas prediktif model dunia secara minimal. Dalam praktiknya, model dunia direduksi menjadi pengode fitur untuk kebijakan, sementara imajinasi, perencanaan, dan optimasi sebagian besar tetap terpisah. Pemisahan ini membatasi generalisasi ke tugas yang lebih menantang dan mencegah kebijakan beradaptasi secara dinamis terhadap sinyal pembelajaran baru. Sebaliknya, kebijakan yang berevolusi sendiri mengintegrasikan sumber umpan balik tambahan dan mekanisme arsitektur, memungkinkan mereka untuk menyesuaikan tujuan dan interaksi mereka dengan model dunia untuk meningkatkan efisiensi, stabilitas, dan generalitas.

Kemajuan terbaru mengarah pada apa yang dapat disebut evolusi diri kebijakan. Tidak seperti pemetaan statis, kebijakan yang berevolusi sendiri menyesuaikan proses pembelajarannya melalui berbagai bentuk panduan. Pendekatan ini dapat dikelompokkan

menjadi tiga kategori utama. Pertama adalah evolusi yang didorong secara eksternal, di mana kebijakan ditingkatkan oleh umpan balik dari pencarian berbasis model atau imajinasi, misalnya dengan menyaring lintasan berkualitas tinggi yang ditemukan melalui perencanaan atau dengan menggabungkan peluncuran yang dibayangkan sebagai masukan kontekstual. Kedua adalah evolusi yang didorong secara internal, di mana kebijakan dioptimalkan secara menyeluruh dalam ruang laten model dunia melalui gradien nilai atau sinyal perbedaan temporal, sehingga memanfaatkan kemampuan diferensiasi model. Ketiga adalah evolusi struktural, yang menekankan modifikasi arsitektur yang memungkinkan kebijakan untuk berkembang dalam panjang rentang dan populasi agen, misalnya melalui transformator modular atau desain multi-agen terdesentralisasi. Subbagian berikut meninjau masing-masing kategori ini secara berurutan dan menyoroti metode yang representatif.

Evolusi yang Didorong Secara Eksternal. Evolusi yang didorong secara eksternal mengintegrasikan perencanaan atau imajinasi ke dalam pembelajaran kebijakan. PlaNet menggabungkan dinamika laten dengan perencana metode entropi silang, dan menyaring urutan aksi yang dihasilkan menjadi sebuah kebijakan, secara efektif mentransfer kompetensi perencanaan ke dalam pengontrol reaktif. MuZero memperluas ide ini dengan menyematkan pencarian pohon Monte Carlo ke dalam model yang dipelajari, secara bersamaan mengoptimalkan kebijakan, nilai, dan fungsi imbalan. Kemajuan teoritis dibuat oleh MPDP, yang memberikan jaminan untuk peningkatan monoton ketika menyaring hasil perencanaan menjadi pembaruan kebijakan. Garis kerja lain, Agen yang Diperkuat Imajinasi (I2A), secara langsung menyediakan peluncuran yang dibayangkan sebagai fitur masukan ke jaringan kebijakan, memungkinkannya untuk belajar bagaimana menggunakan masa depan yang disimulasikan untuk pengambilan keputusan. Metode-metode ini menunjukkan bahwa umpan balik eksternal, baik yang berasal dari pencarian atau imajinasi, dapat mendorong kebijakan melampaui pemetaan statis menuju strategi adaptif dan terinformasi.

Evolusi yang Didorong Secara Internal. Evolusi yang didorong secara internal memanfaatkan tujuan yang dapat dibedakan dalam model dunia untuk secara langsung mengoptimalkan kebijakan. Gradien Nilai yang Dibayangkan (IVG) pertama kali menunjukkan bahwa kebijakan dapat diperbarui dengan menyebarkan gradien melalui lintasan yang dibayangkan di ruang laten. Berdasarkan hal ini, TD-MPC dan TD-MPC2 menggabungkan optimasi lintasan rentang pendek dengan pembelajaran perbedaan temporal, memungkinkan kebijakan untuk menyeimbangkan kualitas prediksi jangka pendek dengan pengembalian rentang panjang. Metode berpengaruh lainnya adalah MBPO, yang menstabilkan pelatihan dengan menggunakan peluncuran sintetis pendek dari model dalam pembelajaran di luar kebijakan, meningkatkan efisiensi dan keandalan. Dalam pendekatan ini, kebijakan berevolusi dari konsumen pasif imbalan menjadi peserta aktif dalam optimasi yang dapat dibedakan, memanfaatkan model dunia sebagai sumber langsung sinyal pembelajaran.

Evolusi Struktural. Seiring lingkungan dan tugas menjadi lebih kompleks, kebijakan harus mampu meningkatkan kapasitas dan mekanisme koordinasinya. Evolusi struktural mengacu pada inovasi arsitektur yang mengadaptasi bentuk kebijakan itu sendiri. UniZero mengusulkan model dunia laten yang terukur berdasarkan transformator modular, yang secara

bersamaan memprediksi dinamika, nilai, dan sinyal kebijakan di berbagai rentang yang lebih luas. Dalam pengaturan multi-agen, MARIE mengembangkan model dunia berbasis transformator terdesentralisasi dengan agregasi terpusat, yang memungkinkan kebijakan untuk berkoordinasi secara efektif di seluruh agen sambil mempertahankan otonomi lokal. Contoh-contoh ini menggambarkan bagaimana evolusi struktural memungkinkan kebijakan untuk beradaptasi tidak hanya melalui sinyal optimasi tetapi juga melalui konfigurasi ulang arsitektur mereka, mendukung penalaran rentang panjang dan kontrol kooperatif di domain yang semakin menuntut.

Evolusi Aktor

Dalam model dunia konvensional, aktor biasanya diimplementasikan sebagai kepala kebijakan sederhana yang memetakan keadaan laten ke dalam tindakan primitif. Desain ini mengasumsikan bahwa model dunia memberikan sinyal prediksi yang cukup akurat, sementara aktor hanya menerjemahkannya ke dalam kontrol yang dapat dieksekusi. Meskipun efektif di lingkungan yang terbatas, aktor yang tidak berevolusi tersebut terbatas kemampuannya untuk melakukan generalisasi di berbagai domain. Mereka kesulitan untuk mengabstraksikan perilaku kompleks, beradaptasi dengan beragam kondisi tugas, atau menyelaraskan diri dengan bias struktural dari model dunia itu sendiri. Akibatnya, aktor konvensional sering menghadapi kesulitan dalam mencapai kontrol yang kuat dalam pengaturan jangka panjang, multimoda, atau dinamis.

Oleh karena itu, riset terbaru telah berupaya untuk mengembangkan modul aktor itu sendiri. Tidak seperti kebijakan statis, aktor yang berevolusi memodifikasi konteks masukan, struktur internal, atau representasi keluaran mereka, memungkinkan mereka untuk beradaptasi bersama dengan dinamika model dunia. Pergeseran ini mencerminkan integrasi yang lebih dalam antara pembangkitan tindakan dan imajinasi. Evolusi masukan memperkaya proses pengambilan keputusan aktor dengan mengkondisikan tindakan pada skenario yang dibayangkan daripada keadaan mentah, menanamkan penalaran berorientasi masa depan langsung ke dalam pemilihan tindakan. Evolusi struktural mengintegrasikan aktor lebih erat dengan alur kerja generatif model dunia, di mana prediksi tindakan dan imajinasi keadaan disatukan dalam satu model. Evolusi keluaran mendefinisikan ulang ruang tindakan dengan mempelajari abstraksi laten atau berbasis lintasan, memungkinkan perilaku yang lebih mudah ditransfer dan terstruktur. Subbagian berikut menguraikan ketiga bentuk evolusi aktor ini dan karya-karya representatifnya.

Konteks Masukan. Salah satu jalur riset memperkaya masukan kepada aktor dengan menggabungkan imajinasi model dunia ke dalam pengambilan keputusan. Alih-alih hanya memetakan dari keadaan laten, aktor mengonsumsi peluncuran imajiner atau ringkasan prediksi masa depan. I2A yang inovatif menunjukkan bahwa aktor dapat belajar menafsirkan dan memanfaatkan prediksi yang dihasilkan oleh model dunia, secara efektif menanamkan wawasan ke depan ke dalam kebijakan mereka. Berdasarkan hal ini, LEXA melatih dua aktor komplementer (penjelajah dan pencapai) sepenuhnya dalam ruang laten imajiner, di mana eksplorasi menemukan beragam keterampilan dan pencapai belajar mencapai tujuan laten. Paradigma ini menunjukkan bahwa dengan mengembangkan konteks masukan, aktor dapat

secara langsung memanfaatkan imajinasi prediktif untuk pengambilan keputusan yang lebih terinformasi dan terarah pada tujuan.

Arsitektur Internal. Aliran riset kedua memikirkan kembali peran struktural aktor dalam model dunia. Alih-alih berfungsi sebagai kepala kebijakan eksternal, aktor tertanam dalam arsitektur generatif model itu sendiri. LaDi-WM mengintegrasikan prediksi tindakan ke dalam kerangka difusi laten, di mana keadaan masa depan dan urutan tindakan disempurnakan secara iteratif dalam ruang laten yang sama. Demikian pula, WorldVLA menyatukan pemodelan visi-bahasa-tindakan dengan model dunia secara autoregresif, secara bersamaan menghasilkan observasi multimoda dan tindakan yang sesuai. Metode-metode ini menggambarkan bagaimana aktor berevolusi secara struktural, menjadi komponen dari alur kerja generatif daripada tambahan, sehingga meningkatkan koherensi antara masa depan yang dibayangkan dan tindakan yang dapat dieksekusi.

Representasi Keluaran. Garis pengembangan lain mendefinisikan ulang bagaimana tindakan direpresentasikan dan dikeluarkan. Alih-alih menghasilkan sinyal kontrol tingkat rendah secara langsung, aktor dapat beroperasi di atas abstraksi tindakan laten. AdaWorld mempelajari representasi aksi dari video tanpa label melalui pelatihan awal mandiri, mengekstrak kode aksi laten yang memfasilitasi adaptasi lintas domain. PreLAR memperluas paradigma ini dengan melatih model dunia pada video tanpa aksi, kemudian memperkenalkan representasi aksi yang dapat dipelajari dalam tugas hilir, menjembatani kesenjangan antara persepsi pasif dan kontrol aktif. Metode representatif lainnya, PolyGRAD, menggunakan model difusi untuk menghasilkan seluruh lintasan, dengan gradien kebijakan yang memandu proses difusi. Hal ini memungkinkan aktor untuk menghasilkan urutan aksi yang konsisten secara temporal, daripada primitif bertahap, meningkatkan efisiensi sampel dan ketahanan perencanaan. Secara kolektif, pendekatan-pendekatan ini menyoroti bagaimana representasi keluaran berevolusi dari perintah motorik langsung menjadi abstraksi laten atau berbasis lintasan yang terstruktur.

Evolusi Imbalan

Dalam model dunia konvensional, fungsi imbalan biasanya diberikan sebagai pengawasan tetap dari lingkungan. Agen bergantung pada sinyal yang didefinisikan secara eksternal ini, yang sering kali dibuat oleh para ahli domain, untuk memandu pembelajaran kebijakan dan estimasi nilai. Desain semacam itu memiliki keterbatasan yang sudah dikenal. Imbalan yang dibuat secara manual memiliki cakupan yang sempit dan mungkin tidak dapat menangkap tujuan tugas yang abstrak atau kompleks. Imbalan seringkali jarang atau tertunda, yang menyebabkan eksplorasi yang tidak efisien dan penugasan kredit yang tidak stabil. Selain itu, sinyal imbalan bersifat statis dan tidak beradaptasi seiring dengan evolusi agen atau model dunia, sehingga membatasi generalisasi di berbagtugas AI dan domain. Dalam pengaturan ini, imbalan tetap menjadi kendala yang tidak fleksibel daripada komponen dinamis yang berinteraksi dengan modul lain.

Pengembangan terbaru telah beralih ke apa yang dapat digambarkan sebagai evolusi diri imbalan, di mana sinyal imbalan tidak lagi tetap tetapi diadaptasi sebagai respons terhadap proses pembelajaran agen dan representasi model dunia. Dibandingkan dengan sinyal statis,

imbalan yang berevolusi sendiri muncul dari struktur internal model, strategi pra-pelatihan yang independen dari tugas, atau keselarasan dengan kriteria evaluasi eksternal. Pendekatan ini menyoroti bahwa imbalan itu sendiri dapat berevolusi sebagai bagian dari sistem pembelajaran. Bagian selanjutnya dari bagian ini meninjau tiga arah representatif: imbalan internal model, pra-pelatihan yang tidak bergantung pada tugas, dan imbalan yang selaras secara eksternal.

Imbalan Internal Model. Kategori pertama memperoleh imbalan langsung dari dinamika dan representasi model dunia. Alih-alih bergantung secara eksklusif pada tujuan yang didefinisikan secara eksternal, sinyal intrinsik muncul dari keadaan laten, kesalahan prediksi, atau perkiraan ketidakpastian. PRIOR Berbasis Tinjau Balik memanfaatkan model maju yang dipelajari untuk menilai pentingnya keadaan individu dalam suatu lintasan dan mendistribusikan kembali pengembalian berbasis preferensi ke transisi kritis, sehingga mengatasi penugasan kredit temporal. LEXA mendefinisikan imbalan dalam imajinasi laten dengan memperkirakan jarak temporal antar keadaan, yang lebih akurat mencerminkan kesulitan mencapai tujuan daripada ukuran tingkat piksel. Metode berorientasi eksplorasi juga termasuk dalam kategori ini. Plan2Explore memperkenalkan perolehan informasi dalam peluncuran laten sebagai sinyal berbasis rasa ingin tahu, Pembelajaran Model Dunia Aktif menggunakan ukuran kemajuan pembelajaran untuk membentuk eksplorasi, dan Melampaui Imajinasi merancang imbalan jangkauan episodik yang mendorong agen untuk memperluas himpunan masa depan yang dapat dijangkau dalam imajinasi. Kontribusi-kontribusi ini mengilustrasikan bagaimana modul imbalan dapat diformulasikan ulang sebagai sinyal adaptif yang berevolusi bersama dengan keadaan internal model dunia.

Pra-pelatihan Tak Bergantung Tugas. Arah kedua menekankan pra-pelatihan tanpa imbalan, di mana model dunia dioptimalkan tanpa akses ke imbalan tugas eksplisit. Paradigma ini menggeser peran imbalan dengan mempersiapkan model untuk memperoleh representasi tujuan umum yang mempercepat pembelajaran berbasis imbalan selanjutnya. CASCADE melatih populasi agen eksplorasi di bawah tujuan teori informasi untuk memaksimalkan keragaman lintasan, sehingga menghasilkan dinamika laten yang dapat digeneralisasikan di berbagai domain. WAKER memajukan ide ini dengan memperkenalkan desain kurikulum untuk pelatihan model dunia tanpa imbalan, memilih contoh lingkungan yang meminimalkan kesalahan prediksi terburuk. Dalam kedua kasus, tidak adanya imbalan selama pra-pelatihan tidak menghambat pembelajaran; sebaliknya, hal itu menghasilkan model dunia dengan representasi yang lebih kuat, yang kemudian memungkinkan adaptasi yang efisien dan stabil setelah imbalan hilir diperkenalkan. Oleh karena itu, pra-pelatihan tak bergantung tugas mewakili bentuk evolusi diri imbalan di mana peran imbalan didefinisikan ulang dari sinyal pelatihan utama menjadi tujuan penyempurnaan.

Imbalan yang Selaras Secara Eksternal. Kategori ketiga mengeksplorasi integrasi metrik yang dapat diverifikasi dan diinterpretasikan ke dalam fungsi imbalan. Tidak seperti metode intrinsik, pendekatan ini menggabungkan kriteria eksternal yang sesuai dengan tujuan yang dapat dipahami manusia. RLVR-World mendemonstrasikan perspektif ini dengan melatih model dunia dengan sinyal imbalan yang dapat diverifikasi, seperti akurasi prediksi atau ukuran

kualitas persepsi. Penyematan metrik tersebut ke dalam proses optimasi memastikan bahwa sinyal imbalan yang dipelajari konsisten dengan standar evaluasi hilir. Meskipun riset ke arah ini masih terbatas, penyelarasan imbalan dengan metrik eksternal membuka jalan yang menjanjikan untuk meningkatkan transparansi, interpretasi, dan keandalan model dunia.

Evolusi Diri Memori

Memori adalah landasan sistem cerdas, yang memungkinkan agen untuk mengakumulasi pengetahuan di berbagai pengalaman, melakukan generalisasi ke situasi baru, dan membangun model lingkungan jangka panjang. Dalam konteks model dunia yang berevolusi sendiri, memori tidak tetap statis tetapi mengalami adaptasi, restrukturisasi, dan abstraksi yang berkelanjutan. Bagian ini memperkenalkan landasan konseptual memori yang berevolusi sendiri, meninjau mekanisme komputasi saat ini, dan menguraikan potensi lintasan untuk merancang arsitektur memori adaptif yang dapat berevolusi seiring dengan tuntutan tugas dan perubahan lingkungan. Memori memainkan peran mendasar dalam membangun dan mengembangkan model dunia. Model dunia, menurut definisinya, bukan hanya representasi statis dari realitas eksternal tetapi substrat kognitif dinamis yang memungkinkan agen untuk memprediksi, bernalar, dan beradaptasi dalam jangka waktu yang panjang. Tanpa mekanisme memori yang efektif, model dunia tetap dangkal, rentan terhadap pergeseran konteks, lupa, dan bahkan halusinasi. Oleh karena itu, mengoptimalkan memori sangat penting untuk memungkinkan model dunia yang berevolusi sendiri yang dapat mengakumulasi pengalaman, menyempurnakan representasi internal, dan memastikan perilaku yang stabil dalam lingkungan yang kompleks dan dinamis.

Tidak seperti metode pelatihan seperti penyempurnaan atau pengeditan pengetahuan, yang secara statis membentuk kembali parameter model, optimasi memori pada waktu inferensi berfokus pada pemanfaatan informasi adaptif selama penalaran. Dalam pandangan ini, model dunia berevolusi tidak hanya berdasarkan apa yang dipelajarinya secara permanen tetapi juga berdasarkan bagaimana ia menjadwalkan, menyimpan, dan mengingat pengetahuan sesuai permintaan. Hal ini menyelaraskan memori dengan pengambilan keputusan: memutuskan apa yang harus diingat, apa yang harus dibuang, dan bagaimana cara menggeneralisasi. Dari perspektif ini, optimasi memori untuk model dunia dapat dikelompokkan menjadi dua proses yang saling melengkapi: (1) *Optimasi memori jangka pendek*, memastikan koherensi lokal dan penalaran yang efisien dalam jendela konteks yang terbatas. (2) *Optimasi memori jangka panjang*, memungkinkan akumulasi dan pengambilan pengalaman yang berkelanjutan di berbagtugas AI dan sesi. Bersama-sama, keduanya menyediakan substrat tempat model dunia berevolusi dari persepsi langsung menjadi kognisi yang berkelanjutan.

Memori Jangka Pendek dalam Model Dunia

Memori jangka pendek berfungsi sebagai “memori kerja” dari sebuah model dunia. Memori ini menyimpan giliran dialog langsung, jejak penalaran perantara, dan isyarat kontekstual yang dibutuhkan untuk menghasilkan respons yang koheren. Mengingat jendela konteks terbatas, strategi seperti peringkasan, retensi selektif, dan perhatian jarang digunakan untuk mencegah kelebihan beban. Peringkasan rekursif secara bertahap memadatkan

pengetahuan menjadi representasi yang ringkas, sementara MemoChat membangun memori tingkat dialog untuk mempertahankan interaksi yang dipersonalisasi. Metode berbasis kompresi seperti COMEDY dan ReadAgent meningkatkan efisiensi konteks dengan menyaring percakapan sebelumnya menjadi jejak tingkat inti. Desain yang lebih canggih mengambil keadaan perantara untuk memandu penalaran multi-langkah, seperti MoT dan StructRAG. Terinspirasi oleh teori kognitif, MemoryBank mengintegrasikan kurva pelupakan Ebbinghaus untuk memperbarui jejak memori secara dinamis berdasarkan kekinian dan relevansi. Refleksi mendorong batas lebih jauh dengan membiarkan agen merefleksikan hasil episodik dan menyimpan pelajaran untuk adaptasi di masa depan. Teknik-teknik ini memperluas stabilitas jangka pendek model dunia, tetapi keterbatasan inherennya terletak pada sifat sementara: mereka tidak dapat mempertahankan evolusi pengetahuan lintas sesi tanpa penyimpanan jangka panjang.

Memori Jangka Panjang dalam Model Dunia

Untuk mencapai evolusi diri, model dunia membutuhkan mekanisme memori jangka panjang yang secara terus-menerus mengakumulasi pengetahuan di luar jendela konteks. Memori jangka panjang memungkinkan agen untuk mengingat pengetahuan faktual, merekonstruksi riwayat tugas, dan menyesuaikan perilaku sesuai dengan interaksi masa lalu, sehingga mengubah pengalaman satu kali menjadi kompetensi yang bertahan lama. Paradigma yang paling berpengaruh adalah Generasi Berbantuan Pengambilan (RAG), di mana model dunia secara aktif menanyakan memori eksternal untuk mendukung penalaran. Sistem seperti EWE dan A-MEM menunjukkan bagaimana memori kerja eksplisit dan pengindeksan dinamis menciptakan jaringan pengetahuan yang berevolusi. Sementara itu, pendekatan pengambilan agen memungkinkan model untuk secara otonom memutuskan apa dan kapan harus mengambil, mengintegrasikan lintasan episodik ke dalam perencanaan masa depan.

Sistem jangka panjang berbasis vektor (misalnya, MemGPT, NeuroCache, G-Memory, AWESOME, Mem0) menyediakan pengkodean memori yang terukur dan akses cepat. Inspirasi biologis, seperti HippoRAG, mensimulasikan mekanisme hipokampus, sementara desain berbasis grafik seperti GraphReader memungkinkan pengambilan terstruktur. Sistem simbolik seperti ChatDB lebih lanjut menempatkan memori dalam basis data eksplisit [26–28]. Di luar penyimpanan, mekanisme kontrol memori mengatur evolusi model dunia. MATTER memilih segmen dari sumber heterogen, sementara AWM mempertahankan pembaruan berkelanjutan di berbagai pengaturan. MemoryBank meninjau kembali pengetahuan lama untuk melawan kelupaan, dan MEM1 menggunakan pembelajaran penguatan untuk menyeimbangkan konsolidasi dan pemangkasan. Sistem seperti MrSteve menambahkan struktur episodik “apa-di mana-kapan”, sementara MIRIX memperkenalkan memori bersama multi-agen untuk penalaran kolaboratif jangka panjang.

Menuju Memori yang Berevolusi Sendiri

Karya-karya terbaru menekankan bahwa memori bukan sekadar unit penyimpanan tetapi substrat yang berevolusi sendiri untuk model dunia. Memori yang efektif memungkinkan agen untuk mengumpulkan, mengatur ulang, dan mengabstraksi pengalaman menjadi heuristik tingkat yang lebih tinggi dan strategi yang dapat digunakan kembali. Expel

menyaring lintasan masa lalu menjadi aturan yang dapat digeneralisasikan, sementara Memori Alur Kerja Agen mencatat alur kerja subtugas yang dapat digunakan kembali. Riche-lieu menunjukkan bagaimana strategi negosiasi meningkat melalui permainan mandiri dan peningkatan memori yang berulang. Mekanisme seperti operasi tambah-gabung-hapus Mem0 atau penandaan semantik MemInsight memastikan koherensi dari waktu ke waktu. INGAT menggabungkan penguatan episodik dengan penyimpanan jangka panjang,

lebih jauh menyoroti evolusi memori untuk evolusi model dunia.Pengoptimalan memori bukanlah peningkatan periferal, melainkan mesin model dunia yang berevolusi sendiri. Memori jangka pendek memastikan koherensi langsung; memori jangka panjang menjamin ketekunan dan adaptasi; dan mekanisme memori yang berevolusi sendiri mengubah pengalaman episodik menjadi kompetensi yang dapat digeneralisasikan. Secara bersama-sama, lapisan-lapisan ini memungkinkan model dunia untuk melampaui pengetahuan statis, terus-menerus merestrukturisasi dirinya sendiri, dan mempertahankan penalaran yang kuat dalam lingkungan yang dinamis.

Evolusi Diri Alat

Alat adalah komponen yang sangat diperlukan dari agen cerdas, yang berfungsi sebagai antarmuka yang melaluinya agen memahami, bertindak, dan memperluas jangkauan mereka ke dunia luar. Mereka menyediakan akses ke database terstruktur, sumber daya komputasi, API, dan layanan khusus domain, yang secara signifikan memperkuat kemampuan pemecahan masalah. Pemanfaatan yang efektif memerlukan agen untuk menentukan kapan, bagaimana, dan mengapa menggunakan alat, dan untuk mengintegrasikan keluaran alat ke dalam penalaran multi-langkah yang koheren. Tantangan ini telah memicu riset ekstensif mengenai pengoptimalan alat dan, yang terbaru, evolusi alat, yang menandai transisi dari penggunaan alat pasif menuju pembuatan dan pengelolaan alat yang otonom. Bagian ini menyatukan upaya yang ada dan menguraikan kemajuan tiga tahap: mengoptimalkan interaksi alat, penguasaan dan penyempurnaan berulang, dan pembuatan otonom dengan manajemen yang dapat diskalakan. Secara keseluruhan, hal-hal ini merupakan landasan evolusi alat adaptif dalam model dunia yang berevolusi dengan sendirinya.

Dari Penggunaan Alat ke Optimasi Alat

Studi awal berfokus pada peningkatan kemampuan agen untuk menggunakan alat yang ada secara efisien. Karena model bahasa besar (LLM) dilatih terutama untuk pembangkitan teks, mereka kurang memiliki pemahaman intrinsik tentang pemanggilan alat. Untuk mengatasi hal ini, para ahli telah mengembangkan strategi optimasi berbasis pelatihan dan inferensi. Optimasi berbasis pelatihan memanfaatkan penyempurnaan terawasi (SFT) pada lintasan penggunaan alat yang dikurasi. Lintasan ini menangkap kueri, langkah penalaran, panggilan alat, dan jawaban, secara eksplisit mengajarkan agen cara merencanakan, mengeksekusi, dan mengintegrasikan alat. Varian seperti ToolLLM dan GPT4Tools menghasilkan data penggunaan alat sintetis, sementara STE dan TOOLEVO mengandalkan eksplorasi simulasi untuk lintasan berkualitas lebih tinggi. Agen multimoda, yang dicontohkan oleh T3-Agen, memperluas teknik ini melampaui teks.

Pembelajaran penguatan (RL) menawarkan jalan lain, memungkinkan agen untuk menyempurnakan kebijakan penggunaan alat melalui interaksi dan umpan balik imbalan, seperti yang terlihat pada ReTool dan Tool-Star. Optimasi waktu inferensi menghindari pelatihan ulang dengan meningkatkan panduan kontekstual selama waktu pengujian. Metode berbasis perintah menyusun ulang dokumentasi alat menjadi instruksi yang ringkas dan ramah LLM, sementara pendekatan berbasis penalaran menggunakan algoritma pencarian pohon atau perencanaan untuk secara dinamis memilih dan menyusun alat. Strategi-strategi ini menunjukkan bahwa bahkan tanpa pembaruan parameter, agen dapat secara signifikan meningkatkan kemampuan penggunaan alat mereka.

Penguasaan Melalui Penyempurnaan Iteratif

Optimasi saja tidak dapat menjamin kekokohan. Alat yang baru dibuat atau jarang digunakan seringkali tetap rapuh dan rawan kesalahan. Oleh karena itu, langkah evolusi kedua melibatkan penyempurnaan dan penguasaan iteratif. Kerangka kerja seperti LearnAct dan Dari Eksplorasi Menuju Penguasaan menyematkan agen dalam lingkaran koreksi diri, di mana kesalahan dari kompilasi, ketidaksesuaian API, atau umpan balik lingkungan dianalisis untuk menemukan penyebab kegagalan. Mengatasi masalah "penugasan kredit" ini, yaitu, mengidentifikasi komponen alat mana yang menyebabkan kegagalan, sangat penting untuk penguasaan yang andal. Penyempurnaan meluas melampaui debugging kode, karena agen juga merevisi dokumentasi alat dan contoh penggunaan, sehingga meningkatkan interpretasi dan penggunaan kembali. Paradigma manusia – didalam - putaran mempercepat proses ini dengan memungkinkan para ahli untuk memvalidasi, mengoreksi, atau memberikan panduan tingkat tinggi. Penyempurnaan hibrida semacam itu tidak hanya meningkatkan keandalan fungsional tetapi juga memastikan keselarasan dengan prinsip keselamatan dan etika. Dalam pengertian ini, penguasaan mengubah skrip alat mentah menjadi kemampuan yang stabil, dapat digunakan kembali, dan dapat dipercaya, memperkuat basis keterampilan model dunia.

Di Luar Penggunaan: Penemuan, Penciptaan, dan Manajemen Otonom

Batas evolusi alat tidak hanya terletak pada penggunaan atau penyempurnaan alat tetapi juga pada penemuan, penciptaan, dan pengelolaan alat secara otonom. Lintasan ini mencerminkan kemajuan peradaban manusia dari penggunaan objek alami sebagai alat hingga rekayasa instrumen yang semakin canggih. Penemuan dan kreasi otonom. Agen seperti Voyager menunjukkan perolehan keterampilan terbuka melalui eksplorasi intrinsik di lingkungan yang kaya seperti Minecraft. Sistem terstruktur seperti CREATOR dan SkillWeaver memformalkan sintesis alat ke dalam API modular, memisahkan kreasi yang dapat digeneralisasikan dari aplikasi khusus domain. Sistem berbantuan pengambilan, yang dicontohkan oleh Alita, secara dinamis menambang atau menghasilkan fungsi setelah menemukan kesenjangan kemampuan, sementara CRAFT merakit perangkat alat khusus domain dari cuplikan kode yang dapat digunakan kembali. Pendekatan-pendekatan ini secara kolektif menunjukkan paradigma yang muncul di mana agen bertindak sebagai pembuat alat daripada sekadar pengguna alat.

Manajemen dan komposisi yang terukur. Seiring bertambahnya pustaka alat, agen harus mengatasi "kutukan kelimpahan". Mekanisme pengambilan sederhana kesulitan ketika

ribuan keterampilan tersedia. Pendekatan seperti ToolGen mengkodekan alat secara langsung ke dalam kosakata model, membingkai ulang pengambilan sebagtugas AI generasi. Sistem tingkat tinggi seperti AgentSquare menggunakan meta pembelajaran di seluruh komponen perencanaan, memori, dan penggunaan alat untuk mengoptimalkan arsitektur modular untuk tugas multi-langkah. Melihat lebih jauh ke depan, Mesin Darwin Gödel membayangkan modifikasi diri rekursif, di mana agen berevolusi tidak hanya perangkat alat mereka tetapi juga arsitektur yang mendasarinya, mengaburkan batas antara "agen" dan "alat."

Menuju Ekosistem Alat yang Berevolusi Sendiri

Perkembangan dari penggunaan alat yang dioptimalkan menuju penciptaan alat otonom dan manajemen yang terukur mewakili lebih dari sekadar kemajuan bertahap; ini menandai pergeseran paradigma menuju ekosistem alat yang berevolusi sendiri. Dalam sistem seperti itu, agen terus-menerus mengidentifikasi kesenjangan kemampuan, menghasilkan dan menyempurnakan alat baru, dan mengelola integrasinya ke dalam repertoar yang koheren. Ini membentuk lingkaran umpan balik tertutup dari penemuan, penguasaan, dan integrasi, yang mencerminkan kemampuan adaptasi rekursif evolusi biologis.

Otonomi juga menghadirkan tantangan baru. Memastikan keamanan, mencegah pembuatan alat yang berbahaya, dan merancang mekanisme verifikasi tetap menjadi isu yang belum terselesaikan. Lebih jauh lagi, mengintegrasikan evolusi alat dengan model dunia yang lebih luas yang mencakup memori, persepsi, dan dinamika membutuhkan kerangka kerja yang terpadu. Terlepas dari tantangan-tantangan ini, lintasannya jelas: masa depan sistem cerdas tidak hanya bergantung pada kemampuan mereka untuk menggunakan alat secara efektif tetapi juga pada kapasitas mereka untuk mengembangkan alat-alat yang memperluas cakrawala kognitif mereka.

Evolusi Diri Pengoptimal Dinamika

Model dunia, cetak biru internal agen tentang lingkungannya, sangat penting untuk mencapai adaptabilitas otonom dalam sistem AI. Dengan mengkodekan keadaan, transisi, dan prediksi, model ini memungkinkan agen untuk mensimulasikan hasil, bernalar melalui tugas multi-langkah, dan beradaptasi dengan konteks dinamis. Pengoptimal dinamika yang berevolusi sendiri menyempurnakan model dunia ini, memastikan model tersebut tetap kuat dan adaptif melalui umpan balik iteratif, pembelajaran, dan integrasi alat. Tidak seperti pengoptimal statis, sistem ini mempertahankan evolusi berkelanjutan dari model dunia, memungkinkan agen untuk mengantisipasi pergeseran lingkungan dan secara proaktif memperbarui pemahaman mereka tanpa intervensi manusia yang ekstensif. Proses ini sangat penting untuk tugas-tugas yang membutuhkan penalaran jangka panjang, seperti perencanaan di domain terbuka atau berinteraksi dengan data waktu nyata.

Optimasi berbasis imbalan adalah landasan untuk mengembangkan model dunia. Agen menyempurnakan representasi internal mereka menggunakan umpan balik lingkungan atau sinyal yang dihasilkan sendiri, memastikan keselarasan dengan dinamika dunia nyata. Misalnya, Tool-Star menggunakan pembelajaran penguatan (RL) untuk memperbarui prediksi model dunia untuk interaksi multi-alat, meningkatkan generalisasi di berbagtugas AI. SPORT memperluas pendekatan ini ke lingkungan multimoda, menerapkan optimasi preferensi

bertahap untuk menyempurnakan representasi visual dan tekstual tanpa anotasi manusia. Metode ini memanfaatkan imbalan tingkat proses untuk menyesuaikan parameter model dunia, berfungsi mirip dengan pembaruan gradien dan memastikan antisipasi yang akurat terhadap keadaan lingkungan. Dalam pengaturan multi-agen, SAGE mengintegrasikan penalaran reflektif dengan optimasi memori yang terinspirasi Ebbinghaus, mengurangi inkonsistensi dalam model dunia selama tugas-tugas kompleks seperti operasi basis data.

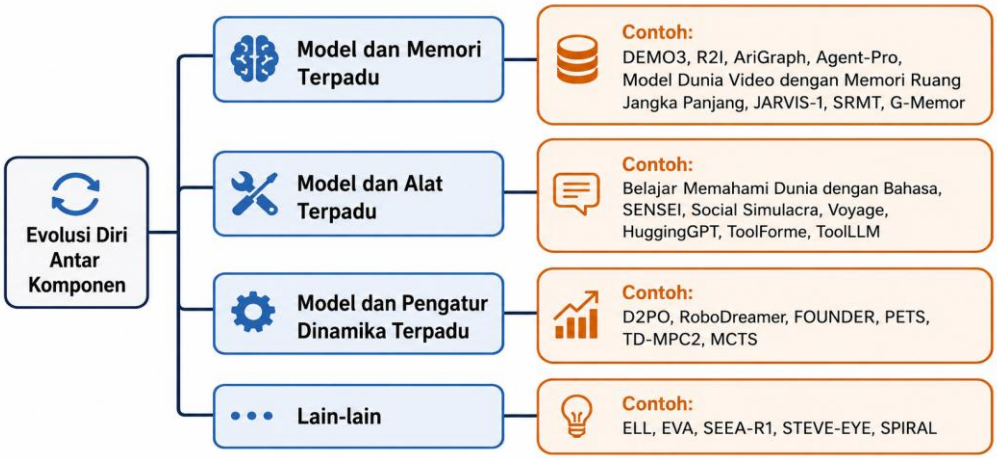
Pembelajaran imitasi dan demonstrasi lebih lanjut memperkuat evolusi model dunia dengan menggabungkan lintasan yang dihasilkan sendiri atau lintas agen. STaR memulai penalaran dengan menyempurnakan model dunia secara iteratif dengan demonstrasi yang berhasil, memungkinkan pembaruan tanpa pengawasan. Pendekatan berbasis populasi, seperti EvoA-gentX, mempertahankan beragam varian model dunia dan memilih konfigurasi optimal melalui kompetisi evolusioner. ToolLLM menghasilkan lintasan penggunaan alat berkualitas tinggi untuk menyempurnakan pemahaman model dunia tentang pemetaan aksi-hasil, terutama dalam skenario multi-giliran. Gorilla meningkatkan kemampuan adaptasi dengan mengintegrasikan pengambil dokumen, mendasarkan prediksi model dunia pada pengetahuan yang diambil saat alat baru diperkenalkan. Adaptasi khusus domain menyoroti peran pengoptimal dalam menyesuaikan model dunia dengan konteks khusus. Dalam tugas pengkodean dan GUI, Skywork-Reward-V2 mengkurasi data preferensi melalui kolaborasi manusia-AI, menyempurnakan representasi khusus domain untuk presisi prediksi yang lebih besar. ToolGen mengkodekan alat sebagai unit dalam model dunia, menyederhanakan pemilihan dan komposisi untuk tugas skala besar. Evaluasi pada AgentBench melaporkan peningkatan 5–48% pada model sumber terbuka, yang menunjukkan kemampuan pengoptimal untuk meningkatkan daya prediksi.

Meskipun demikian, tantangan tetap ada dalam memastikan keamanan, mencegah kelupaan katastrofik, dan penskalaan ke berbagai domain, yang membutuhkan metrik adaptivitas yang kuat. Arah masa depan mencakup eksplorasi sistem multi-agen ko-evolusioner untuk menyempurnakan model dunia bersama, sehingga meningkatkan generalisasi. Dengan mendorong proses putaran tertutup yang mengintegrasikan umpan balik, peningkatan memori, dan pembaruan yang digerakkan oleh alat, pengoptimal dinamika yang berevolusi sendiri memberdayakan model dunia untuk maju menuju otonomi yang hampir supercerdas. Sistem seperti itu akan semakin mampu menavigasi lingkungan yang kompleks dan selalu berubah dengan ketahanan dan pandangan ke depan.

5.4 EVOLUSI DIRI ANTAR-KOMPONEN

Transisi dari agen AI statis ke agen yang berevolusi sendiri memerlukan inovasi arsitektur yang memungkinkan peningkatan otonom dan berkelanjutan. Evolusi ini tidak monolitik tetapi terjadi di seluruh komponen inti agen, modelnya, memori, alat, dan pengoptimal dinamika, baik secara individual maupun bersama-sama. Bagian ini menguraikan paradigma arsitektur utama yang memfasilitasi pertumbuhan mandiri ini, mengkategorikannya berdasarkan komponen yang terlibat dalam siklus evolusi. Untuk

mengilustrasikan metode ini dan lintasan pengembangannya secara lebih intuitif, kami menyediakan diagram ikhtisar pada Gambar 5.5 yang berfokus pada setiap komponen inti.



Gambar 5.5 Gambaran umum evolusi diri antar-komponen

Evolusi Diri Bersama Model-Memori

Model dirancang untuk menyediakan agen cerdas dengan simulator internal dinamika lingkungan, sehingga mendukung kemampuan inti seperti prediksi, perencanaan, dan pengambilan keputusan. Pada saat yang sama, sistem memori menyimpan dan mengambil informasi di berbagai skala temporal, termasuk memori jangka pendek (STM) untuk penalaran langsung dan memori jangka panjang (LTM) untuk abstraksi pengetahuan. Dalam bab ini, kita mendefinisikan ko-evolusi model dan memori sebagai berikut: sementara parameter dan struktur model dunia terus diperbarui, sistem memori juga diperbarui dan diperluas secara dinamis; melalui interaksi keduanya, terbentuklah sebuah lingkaran tertutup, sehingga memungkinkan pembelajaran jangka panjang adaptif dan generalisasi lintas tugas. Kerangka kerja ko-evolusi model-memori yang khas dapat digambarkan sebagai proses interaktif siklik: model melakukan prediksi dan perencanaan dengan dukungan memori; pengalaman kritis yang dihasilkan selama interaksi secara selektif dituliskan ke dalam dan/atau diorganisasi ulang; kebijakan terus dioptimalkan di bawah pengambilan dan pemutaran ulang memori; dan akhirnya, parameter model diperbarui. Mekanisme lingkaran tertutup ini mencirikan dinamika evolusi diri dari "prediksi model-pembaruan memori-pengambilan keputusan tindakan-pembaruan model," yang menyediakan landasan teoretis terpadu untuk ko-evolusi berkelanjutan antara model dan memori.

Pentingnya ko-evolusi model-memori terletak pada memastikan bahwa model dunia dapat secara dinamis memodelkan berbagai skala temporal. Pada skala waktu cepat, memori jangka pendek diperbarui hampir secara waktu nyata, berfungsi untuk melengkapi pengamatan yang hilang dan mempertahankan konteks lokal. Pada skala waktu menengah, model terus berevolusi melalui pemutaran ulang memori dan pengambilan sampel, mengurangi kelupaan katastrofik. Pada skala waktu lambat, konsolidasi memori jangka panjang dan internalisasi parameter model mendorong abstraksi pengetahuan dan transfer lintas tugas. Integrasi multi-level ini memungkinkan model dunia untuk beradaptasi dengan

cepat terhadap perubahan lingkungan dan secara bertahap mengakumulasi pengetahuan jangka panjang yang stabil.

Dalam riset yang ada, ko-evolusi model dan memori umumnya dapat dibagi menjadi tiga mekanisme yang berbeda. Pertama, paradigma pertama menekankan peran penggerak langsung memori dalam evolusi model. Dalam paradigma ini, memori dianggap sebagai pengalaman eksternal atau sumber pengetahuan, yang digunakan untuk memandu pembaruan parameter model dunia dan optimasi kebijakan. Misalnya, DEMO3 memperkenalkan lintasan demonstrasi sebagai memori eksplisit, mengubah sinyal imbalan jarang menjadi umpan balik yang dapat dipelajari, dan mewujudkan optimasi bersama memori dan model dunia. Metode jenis ini menunjukkan bahwa memori bukan hanya penyimpanan statis, tetapi juga dapat diubah menjadi sinyal pelatihan yang efektif dalam bentuk pemutaran ulang dan penggunaan kembali. Demikian pula, R2I menunjukkan bagaimana menggunakan pengambilan memori untuk menghasilkan masukan yang divisualisasikan atau memutar ulang fragmen kunci, sehingga meningkatkan kinerja model dalam prediksi dan penalaran jangka panjang. Sebaliknya, RAG mewujudkan desain "memori dalam lingkaran". Ia tidak memperbarui parameter model, tetapi dapat meningkatkan penalaran melalui memori nonparametrik eksternal, memungkinkan model untuk mengakses pengetahuan faktual secara adaptif selama inferensi.

Paradigma lain menyoroti pembaruan dan reorganisasi memori secara aktif oleh model. Dalam hal ini, model tidak hanya mengambil memori, tetapi juga dapat menyusun dan mengabstraksikannya melalui penalaran dan prediksi. Salah satu contoh karya yang umum adalah AriGraph, yang dalam tugas simulasi yang diwujudkan, mengkodekan pengalaman ke dalam grafik fakta dan grafik peristiwa yang terstruktur, sehingga membentuk representasi memori yang terstruktur untuk mendukung transfer tugas dan penalaran yang kompleks. Pada saat yang sama, Agent-Pro mengusulkan metode untuk mengubah riwayat interaksi menjadi “keyakinan” dan menyesuaikan perilaku melalui mekanisme refleksi tingkat kebijakan. Proses ini pada dasarnya adalah rekonstruksi dan abstraksi memori yang digerakkan oleh model, yang menyoroti nilai memori dalam pengambilan keputusan jangka panjang.

Paradigma ketiga berfokus pada ko-evolusi putaran tertutup, di mana model dan memori terus menerus saling memperbarui melalui siklus prediksi, penyimpanan, pengambilan, dan pembaruan. Dalam skenario agen tunggal, Model Dunia Video dengan Memori Spasial Jangka Panjang menyematkan memori spasial ke dalam arsitektur model dunia, memungkinkan keuntungan dalam konsistensi adegan dan penalaran jangka panjang. JARVIS-1 menggabungkan memori multimoda dengan agen yang berpusat pada model bahasa, meningkatkan kemampuan perencanaan dan eksekusi dalam tugas dunia terbuka. Dalam skenario multi-agen, mekanisme serupa juga sangat penting. SRMT mengusulkan arsitektur ruang kerja bersama, yang memungkinkan beberapa agen untuk menggabungkan pengamatan lokal dan hasil penalaran, dan kemudian berkoordinasi melalui memori global, sehingga mencapai solusi yang efisien untuk tugas jangka panjang. G-Memory lebih lanjut mengembangkan struktur memori tiga tingkat hierarkis (grafik wawasan, kueri, dan interaksi),

yang mekanisme evolusi berkelanjutannya dengan pengalaman secara signifikan meningkatkan kinerja sistem multi-agen dalam penalaran yang kompleks.

Evolusi Diri Bersama Model-Alat

Evolusi bersama model-alat bertujuan untuk meningkatkan kemampuan pemodelan dunia dari agen terwujud melalui pengenalan alat eksternal, sehingga memungkinkan fungsi inti yang lebih kuat seperti prediksi, perencanaan, dan pengambilan keputusan. Alat menyediakan antarmuka kemampuan multidimensi untuk agen, termasuk persepsi, pengetahuan dan pengambilan, penalaran dan perencanaan, serta eksekusi, yang memungkinkan model untuk menyelesaikan tugas secara lebih efektif di lingkungan yang kompleks dan terbuka. Dalam bab ini, kami mendefinisikan evolusi bersama model-alat sebagai berikut: sementara parameter dan struktur model dunia terus diperbarui, pemilihan, pemanggilan, dan komposisi alat juga berevolusi dengan pengalaman yang terakumulasi; keduanya berinteraksi untuk membentuk lingkaran tertutup. Kerangka kerja evolusi bersama model-alat yang khas dapat digambarkan sebagai proses interaksi siklik: model melakukan penalaran dan perencanaan berdasarkan tujuan tugas, dan sesuai dengan itu memilih dan memanggil alat yang tepat untuk mendapatkan dukungan dalam persepsi, pengetahuan, komputasi, atau eksekusi; Hasil yang dikembalikan oleh alat, sebagai data eksternal yang dapat diverifikasi, dimasukkan ke dalam penyempurnaan keputusan selanjutnya; dengan akumulasi pengalaman interaksi, sistem di satu sisi memperbarui parameter dan strategi model, dan di sisi lain juga memperbarui strategi pada lapisan alat.

Pentingnya ko-evolusi model-alat terletak pertama pada kemampuannya untuk menjembatani kesenjangan kemampuan model berparameter: dengan menggunakan persepsi khusus dan basis pengetahuan eksternal, model dapat mencapai batas kemampuan yang dapat diperluas tanpa memperbesar skala parameter; kedua, keluaran alat (seperti hasil kode yang dapat dieksekusi, bukti mentah yang diambil, atau umpan balik dari eksekusi fisik) menyediakan rantai bukti yang dapat diverifikasi dan dilacak untuk penalaran dan pengambilan keputusan, sehingga secara signifikan meningkatkan keandalan dan interpretasi sistem; akhirnya, alat, sebagai "modul pasang dan pakai" eksternal, secara alami mendukung komposisi dan transferabilitas, di mana rangkaian alat yang sama dapat digunakan kembali di berbagtugas AI, memungkinkan agen untuk memiliki kemampuan generalisasi lintas domain dan adaptasi sedikit contoh.

Dalam studi ko-evolusi model-alat, alat secara umum dapat dibagi menjadi tiga kategori. Alat persepsi terutama bertanggung jawab untuk memperoleh dan merepresentasikan informasi lingkungan. Misalnya, dalam SENSEI, model dasar digunakan sebagai alat persepsi untuk penilaian semantik dan ekstraksi isyarat. Keluarannya langsung dimasukkan ke dalam model dunia dan putaran pembelajaran kebijakan, berfungsi sebagai sumber panduan eksplorasi atau imbalan intrinsik; model kemudian menyesuaikan pemilihan observasi dan distribusi tindakan sesuai dengan itu, yang pada gilirannya mengubah strategi kueri alat. Alat-alat tersebut memperluas cakupan persepsi model dalam lingkungan yang kompleks, memungkinkan agen untuk menangkap dan memanfaatkan informasi semantik dengan lebih efisien.

Sebaliknya, alat penalaran dan perencanaan menekankan penyediaan inferensi logis dan dukungan pengambilan keputusan untuk agen dalam tugas-tugas kompleks. Sebuah karya representatif adalah Belajar Memodelkan Dunia dengan Bahasa, di mana bahasa dianggap sebagai pembawa pengetahuan eksplisit dan kerangka simbolik; selama penalaran, model mengambil atau secara kondisional menghasilkan aturan dan fakta yang relevan untuk mengkompensasi keterbatasan pengetahuan yang diparameterisasi. Selain itu, simulasi sosial membangun cermin virtual dari sistem komputasi sosial untuk menghasilkan dan menyimpan beragam fakta dan pengalaman sosial. Model ini menguji dan merevisi hipotesis dalam masyarakat yang disimulasikan, dan "fakta sosial" yang baru dihasilkan ditulis kembali ke dalam basis pengetahuan untuk mendorong putaran pengambilan dan pengambilan keputusan berikutnya.

Alat eksekusi membawa komputasi eksternal dan tindakan dunia nyata ke dalam lingkaran tertutup yang dapat diverifikasi, yang mewakili skenario paling intuitif dari ko-evolusi model-alat. Voyager memanfaatkan eksekusi kode dan pustaka keterampilan untuk perluasan diri yang berkelanjutan: model mensintesis atau memanggil program berdasarkan dekomposisi tugas, dan hasil eksekusi diumpankan kembali sebagai pengawasan yang kuat dan bukti yang dapat dilacak, mendorong pembaruan dua arah kebijakan dan pustaka keterampilan (termasuk penambahan, rekonstruksi, dan penghapusan keterampilan). Keterampilan yang stabil dan efektif dapat diabstraksikan ke dalam alat yang dapat digunakan kembali tingkat yang lebih tinggi, yang kemudian digabungkan dalam tugas-tugas selanjutnya. Hugging-GPT menggunakan meta-penjadwalan untuk mengarahkan alat di antara beberapa model eksternal, dengan keluaran eksekusi secara langsung mengukur penyelesaian tugas dan sumber kesalahan, sehingga memandu penyesuaian strategi perutean dan penugasan subtugas pada iterasi berikutnya. ToolFormer dan ToolLLM belajar selama pelatihan "kapan dan bagaimana memanggil API," menggeser keputusan pemanggilan dari heuristik ke strategi berbasis data; hasil eksekusi memberikan sinyal pengawasan yang dapat diverifikasi untuk model, membentuk siklus penguatan diri "pemanggilan-verifikasi-internalisasi-pemanggilan ulang."

Evolusi Diri Bersama Model-Pengoptimal Dinamika

Evolusi bersama pengoptimal model-dinamika bertujuan untuk meningkatkan prediktabilitas dinamika lingkungan dan kualitas pengambilan keputusan melalui evolusi terkoordinasi dari parameter model dan mekanisme optimasi. Dalam bab ini, kami mendefinisikan evolusi bersama pengoptimal model-dinamika sebagai berikut: selama proses pembelajaran agen, sementara parameter dan struktur model terus berevolusi, strategi dan mekanisme pembaruan pengoptimal juga terus diadaptasi; keduanya berinteraksi untuk membentuk sebuah lingkaran, sehingga mencapai konvergensi yang lebih cepat dan ketahanan yang lebih kuat. Kerangka kerja bersama yang khas dapat digambarkan sebagai interaksi siklik: model dunia melakukan prediksi dan perencanaan di bawah konfigurasi imbalan dan kebijakan saat ini, mengumpulkan pengembalian lingkungan dan kesalahan model; pengoptimal, berdasarkan hal ini, mengintegrasikan imbalan eksternal dan keluaran dari model imbalan yang dipelajari untuk menghasilkan fungsi tujuan dan jadwal kebijakan.

Selanjutnya, sistem memperbarui model dunia dan parameter kebijakan menggunakan fungsi tujuan ini, sementara pada saat yang sama melatih ulang dan merekonstruksi model imbalan dan aturan penjadwalan kebijakan. Pengoptimal terus-menerus menulis ulang tujuan, dan model, pada gilirannya, mempelajari dinamika dan pengambilan keputusan yang lebih kuat, yang selanjutnya mendorong pengoptimal untuk konvergen menuju konfigurasi imbalan dan kebijakan yang lebih efektif. Nilai inti dari ko-evolusi model-pengoptimal terletak pada koreksi dan penguatan sinyal pembelajaran, sehingga mencapai pembelajaran yang stabil dan efisien dalam lingkungan jangka panjang dan imbalan yang jarang. Pertama, ko-evolusi bersama dapat mengurangi kelangkaan imbalan dengan mengubah tujuan yang sulit dipelajari menjadi sinyal padat yang dapat dipelajari dan secara adaptif mengubah skala besaran dan bobot di seluruh lingkungan dan tahapan; kedua, melalui lingkaran tertutup "imbalan-kebijakan-model," sistem dapat secara bersamaan meningkatkan efisiensi sampel dan generalisasi, memungkinkan model untuk konvergen lebih cepat, sementara prediksi yang stabil memungkinkan pengoptimal untuk mengalokasikan kredit dan merancang sub-tujuan dengan akurasi yang lebih besar.

Dalam riset yang ada, satu kelas metode menekankan ko-evolusi berbasis imbalan. Isu inti dari kelas metode ini terletak pada kenyataan bahwa pengoptimal, melalui mekanisme seperti pemodelan preferensi dan pembaruan imbalan, bertindak langsung pada fungsi kerugian, sehingga mengubah lintasan pembaruan model dunia dan kebijakan. Karya-karya representatif meliputi D2PO, yang memanfaatkan preferensi manusia untuk memberi peringkat lintasan yang dibayangkan, memungkinkan model imbalan, model dunia, dan kebijakan dioptimalkan bersama di bawah tujuan yang terpadu; RoboDreamer mendasarkan tujuan bahasa tingkat tinggi ke dalam unit tindakan tingkat rendah dan memilih urutan tindakan sesuai dengan probabilitas keberhasilan yang diprediksi, pada dasarnya memperlakukan probabilitas keberhasilan sebagai tujuan optimasi untuk memandu generasi dan perencanaan; FOUNDER memperkenalkan mekanisme penyelarasan semantik model dasar ke dalam model dunia dan putaran pengambilan keputusan, sehingga niat tugas abstrak diubah menjadi imbalan dan batasan yang dapat dioperasikan.

Kelas metode lain menekankan ko-evolusi berbasis pencarian. Inti dari kelas ini terletak pada pengoptimal yang mengadopsi optimasi lintasan, strategi pengambilan sampel, dan pencarian pohon sebagai bentuk utamanya, secara langsung memengaruhi pembangkitan dan pemilihan urutan kandidat, dan memberikan lintasan yang diperoleh dari pencarian kembali sebagai sinyal pelatihan ke model dunia dan kebijakan. Metode PETS yang diusulkan oleh Chua dan kawan kawan menggunakan dinamika ansambel dan optimasi berbasis pengambilan sampel (CEM/MPPI), dan menggunakan sinyal ketidakpastian untuk mengatur proses pengambilan sampel dan pembaruan kebijakan; TD-MPC2 lebih lanjut menanamkan pengoptimal lintasan ke dalam ruang laten model dunia, menjadikan hasil perencanaan secara langsung sebagai bagian dari tujuan pelatihan; dan MCTS bersama dengan variannya melakukan pencarian pohon pada model dunia untuk mengevaluasi tindakan kandidat secara prospektif, dan menggunakan distribusi dan pengembalian pencarian yang dihasilkan sebagai sinyal pengawasan untuk memperbarui kebijakan dan fungsi nilai.

Lainnya

Dalam hasil pengamatan yang ada, terdapat juga studi yang mengeksplorasi evolusi diri bersama dari memori dan pengoptimal. ELL berfokus pada kerangka pembelajaran sepanjang hayat yang didorong oleh pengalaman jangka panjang. Metode ini menyimpan dan memutar ulang pengalaman dari proses interaksi dan langsung menggunakannya sebagai sinyal masukan ke pengoptimal, untuk tujuan membentuk imbalan dan mengkalibrasi ulang strategi, sehingga mencapai keterkaitan yang erat antara memori dan pengoptimal dinamis. Metode tersebut secara dinamis menyesuaikan tujuan pembelajaran dan strategi optimasi melalui akumulasi dan pemanggilan pengalaman.

Pada saat yang sama, jalur riset lain menunjukkan mekanisme yang lebih kompleks dari optimasi bersama multi-faktor (Model + Memori + Alat + Pengoptimal). Karya representatif termasuk EVA, yang ide intinya adalah memperkenalkan mekanisme refleksi yang memperlakukan kesalahan prediksi dan kasus kegagalan sebagai pengalaman yang diingat, menghasilkan sinyal korektif melalui modul refleksi eksternal, dan dengan demikian mendorong revisi tujuan pengoptimal sambil mengoordinasikan pembaruan model prediksi video. Demikian pula, SEEA-R1 mengusulkan metode penyempurnaan penguatan terstruktur pohon, di mana penyimpanan pengalaman bercabang berperan sebagai memori, proses penalaran terstruktur pohon berfungsi sebagai alat untuk pengambilan keputusan, dan pengoptimal mencapai evolusi berkelanjutan dari model dan kebijakan melalui penjadwalan hierarkis. Selain itu, STEVE-EYE menggabungkan LLM dengan alat persepsi visual untuk menyediakan pemahaman visual kepada agen terwujud di lingkungan terbuka. Modul persepsinya tidak hanya bertindak sebagai antarmuka alat untuk menyediakan aliran informasi eksternal, tetapi juga mengakumulasi pengalaman persepsi sebagai memori; sinyal-sinyal ini dimasukkan ke dalam putaran optimasi imbalan dan kebijakan, sehingga mendorong model untuk terus berevolusi dalam siklus persepsi-keputusan-optimasi.

Terakhir, SPIRAL mengilustrasikan mekanisme koordinasi multi-faktor dalam lingkungan permainan mandiri multi-agen. Metode ini bergantung pada riwayat interaksi sebagai sumber daya memori, sambil memanfaatkan penalaran multi-giliran dan permainan sebagai alat eksternal, dan pada akhirnya, di bawah pembentukan pengoptimal yang bersifat antagonis, meningkatkan kemampuan penalaran dan kapasitas pemilihan strategi model. Secara keseluruhan, lebih banyak pengembangan yang berupaya mengintegrasikan model, memori, alat, dan pengoptimal secara mendalam. Dalam kerangka kerja ini, pengoptimal tidak lagi berfungsi secara terisolasi sebagai pembaruan parameter, tetapi membentuk lingkaran evolusi diri bersama dengan memori dan alat, memberikan ide-ide baru bagi agen terwujud untuk mencapai adaptasi jangka panjang dan transfer lintas tugas di lingkungan yang kompleks.

BAB 6
AGEN YANG BEREVOLUSI SENDIRI

6.1 PENDAHULUAN

Bab ini menyajikan tinjauan sistematis dan eksplorasi metodologis tentang Agen yang Berevolusi Sendiri. Pertama, pada bagian tinjauan umum, bab ini memperkenalkan latar belakang pengembangan dan motivasi utama dari agen yang berevolusi sendiri. Bab ini kemudian memberikan analisis mendalam tentang Komponen Kunci mereka, termasuk model, memori, dan alat, yang menjelaskan peran masing-masing dalam proses evolusi. Berdasarkan hal ini, bab ini merangkum arsitektur utama, meneliti evolusi diri model, evolusi gabungan model-memori, evolusi gabungan model-alat, dan evolusi gabungan model-memori-alat, sehingga menunjukkan potensi optimasi kolaboratif multidimensi.

Lebih lanjut, bab ini menguraikan strategi utama, yang meliputi evolusi diri berbasis imitasi dan evolusi diri berbasis interaksi, menyoroti mode pembelajaran dan adaptasi yang berbeda dalam mekanisme ini. Pada bagian studi kasus, bab ini menyajikan kerangka kerja praktis lengkap untuk agen yang berevolusi sendiri. Secara keseluruhan, bab ini tidak hanya mensistematiskan dasar-dasar teoritis dan teknis dari agen yang berevolusi sendiri, tetapi juga mengilustrasikan potensi aplikasinya melalui contoh-contoh konkret, menawarkan kerangka kerja referensi terstruktur untuk riset dan praktik di masa mendatang. Agen berbasis Model Bahasa Besar (LLM) telah membuat kemajuan signifikan dalam perencanaan, penalaran, dan pemecahan masalah berbasis alat. Namun, setelah diterapkan, agen-agen ini tetap statis: parameter dan strukturnya tetap. Hal ini membatasi kemampuan mereka untuk menangani tugas-tugas baru, beradaptasi dengan domain yang berubah, atau menanggapi kebutuhan pengguna yang berkembang. Akibatnya, otonomi dan keandalan jangka panjang mereka di lingkungan dinamis tetap terbatas. Untuk mengatasi keterbatasan ini, riset terbaru mulai berfokus pada agen yang berevolusi sendiri, yang terus meningkatkan kemampuan persepsi, penalaran, dan interaksi mereka selama operasi.

Melalui pembelajaran berkelanjutan, adaptasi modular, dan plastisitas arsitektur, agen-agen ini bertujuan untuk tetap tangguh dalam kondisi yang berubah sambil menghindari kelupaan katastrofik. Agen yang berevolusi sendiri biasanya terdiri dari tiga komponen inti: model, memori, dan alat. Model, yang biasanya berupa model bahasa besar atau model multimoda besar, menyediakan kemampuan dalam penalaran semantik, abstraksi, dan pengambilan keputusan. Memori mendukung akumulasi pengetahuan jangka panjang, landasan kontekstual, dan peningkatan kinerja melalui integrasi pengalaman masa lalu. Alat memperluas kemampuan agen untuk berinteraksi dengan lingkungan digital dan fisik, memungkinkan akuisisi informasi, eksekusi tindakan, dan koordinasi alur kerja yang kompleks. Masing-masing komponen ini dapat berevolusi: model beradaptasi melalui pelatihan awal yang berkelanjutan, penyempurnaan, atau pembaruan parameter yang efisien ; memori diatur

ulang untuk mendukung pengambilan yang terukur dan pembelajaran sepanjang hayat; dan alat berkembang secara dinamis untuk menangani tugas atau domain baru.

Meskipun setiap komponen telah berkembang secara independen, tantangan utama selanjutnya adalah mengembangkan arsitektur di mana model, memori, dan alat berevolusi bersama. Kerangka kerja bersama mendorong sinergi yang lebih dalam di antara subsistem, memungkinkan mereka untuk beradaptasi bersama dengan cara yang saling memperkuat. Misalnya, peningkatan dalam organisasi memori dapat memberikan pengetahuan terstruktur untuk memandu adaptasi model, sementara alat baru dapat memperluas ruang aksi model dan membentuk kembali cara memori menyimpan dan mengambil informasi. Karya terbaru seperti EvoAgent mengilustrasikan pendekatan ini dengan mengintegrasikan penalaran multimoda dengan model dunia dan penggunaan alat adaptif, memungkinkan perencanaan mandiri, refleksi diri, dan pengendalian diri secara otonom di berbagai lingkungan. Demikian pula, arsitektur berbasis prototipe modular seperti MoDP menunjukkan bagaimana pemisahan dan komposisi dapat mendukung peningkatan diri secara bertahap pada agen yang terwujud.

Oleh karena itu, mengembangkan mekanisme yang memungkinkan evolusi bersama model, memori, dan alat sangat penting untuk membangun sistem yang benar-benar otonom dan terus menerus meningkatkan diri. Agen semacam itu akan melampaui keterbatasan pendekatan berbasis LLM statis dan membangun fondasi bagi kecerdasan yang menunjukkan pertumbuhan berkelanjutan, pembelajaran terbuka, dan operasi yang andal dalam lingkungan dunia nyata yang dinamis. Pada akhirnya, tujuan agen yang berevolusi sendiri bukan hanya untuk meningkatkan sistem AI saat ini secara teknis, tetapi juga untuk membuka jalan bagi paradigma baru kecerdasan buatan yang adaptif, tangguh, dan belajar sepanjang hayat.

6.2 KOMPONEN UTAMA AGEN YANG BEREVOLUSI SENDIRI

Konstruksi agen yang berevolusi sendiri bergantung pada serangkaian komponen yang saling bergantung yang secara bersama-sama memungkinkan adaptasi, refleksi, dan pertumbuhan jangka panjang. Tidak seperti sistem konvensional di mana penalaran, tindakan, dan memori sering diperlakukan sebagai fungsi yang terisolasi, agen yang berevolusi sendiri membutuhkan desain yang terhubung erat di mana model, memori, dan alat membentuk lingkaran umpan balik yang berkelanjutan. Model menyediakan substrat kognitif untuk penalaran. Memori berfungsi sebagai substrat untuk landasan langsung dan repositori untuk pengalaman sepanjang hayat, dan alat memperluas jangkauan agen melampaui kapasitas intrinsiknya. Secara bersama-sama, komponen-komponen ini mendefinisikan mekanisme di mana agen tidak hanya dapat bereaksi terhadap lingkungan tetapi juga secara iteratif meningkatkan diri mereka sendiri, menjadikan evolusi diri sebagai proses yang berkelanjutan dan bukan desain statis.

Model

Model berfungsi sebagai inti kognitif dari agen yang berevolusi sendiri, menentukan bagaimana ia menafsirkan tugas, menghasilkan strategi, dan menghasilkan respons. Dalam sebagian besar sistem, peran ini dilakukan oleh model dasar seperti model bahasa besar, yang

menawarkan kemampuan linguistik dan penalaran yang luas. Meskipun model-model ini berkinerja baik pada tolok ukur statis, agen yang berevolusi sendiri membutuhkan model-model tersebut untuk melampaui penalaran tetap dan mendukung proses adaptif yang memungkinkan peningkatan berkelanjutan. Studi terbaru menyoroti bahwa model-model tersebut harus menggabungkan metakognisi dan mekanisme penilaian diri, yang memungkinkan mereka untuk memantau penalaran mereka sendiri, mendeteksi kesalahan, dan memperbaiki strategi dari waktu ke waktu. Hal ini membedakan mereka dari model konvensional yang hanya bergantung pada pelatihan awal dan penyempurnaan, karena model yang berevolusi sendiri memperoleh kemampuan untuk meningkatkan diri secara otonom melalui interaksi, pembangkitan data yang diawasi sendiri, dan umpan balik iteratif.

Sejumlah besar riset menunjukkan bahwa agen dapat menghasilkan sinyal pengawasan mereka sendiri untuk meningkatkan model mereka. Misalnya, agen dapat menciptakan masalah, mencoba solusi, dan menggunakan hasilnya sebagai data pelatihan, memungkinkan mereka untuk belajar tanpa anotasi manusia. Ide ini merupakan inti dari metode seperti permainan mandiri, mekanisme pemberian imbalan mandiri, dan penyempurnaan reflektif, yang semuanya membantu model meningkatkan kemampuan penalaran dan perencanaannya. Lebih lanjut, umpan balik interaktif dari lingkungan atau agen lain dapat digunakan sebagai sinyal imbalan, yang mengarah pada peningkatan kualitas kebijakan dan penalaran secara stabil. Bersama-sama, kemajuan ini menunjukkan bahwa evolusi model melibatkan lebih dari sekadar penyesuaian parameter—ini membutuhkan penyematan agen dalam siklus eksplorasi, deteksi kesalahan, dan koreksi yang berkelanjutan untuk mendukung pertumbuhan jangka panjang.

Kebijakan

Kebijakan adalah lapisan operasional dari agen yang berevolusi sendiri, yang mengontrol bagaimana ia memilih dan mengeksekusi tindakan. Tidak seperti agen tradisional dengan kebijakan tetap yang dibuat secara manual, agen yang berevolusi sendiri menyesuaikan strategi mereka berdasarkan pengalaman, mencerminkan gagasan bahwa kebijakan harus berevolusi melalui umpan balik. Adaptasi ini dapat dicapai menggunakan pembelajaran penguatan, penyesuaian reflektif, atau metode hibrida yang menggabungkan imitasi dengan optimasi berbasis hasil. Kemampuan kunci adalah menghasilkan atau mengumpulkan umpan balik tanpa masukan manusia, memungkinkan kebijakan untuk terus diperbarui saat agen beroperasi di lingkungan yang kompleks dan dinamis.

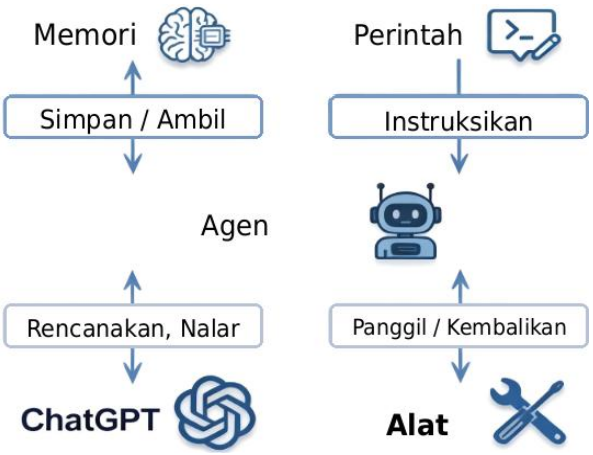
Riset terbaru telah mengeksplorasi metode yang memungkinkan agen untuk menguji kebijakan mereka sendiri dengan membuat kasus uji, mengevaluasi kinerja, dan merevisi strategi sesuai dengan itu. Metode bermain sendiri, di mana agen bersaing melawan versi diri mereka sebelumnya, menciptakan tekanan evolusioner yang memperkuat strategi dari waktu ke waktu. Pendekatan lain menggunakan penyempurnaan yang bersifat antagonis, di mana agen meningkatkan diri dengan mengatasi tantangan yang semakin sulit atau menipu yang dihasilkan secara internal. Metode-metode ini memiliki ide yang sama: pengalaman dan umpan balik agen sendiri berfungsi sebagai sumber utama pembelajaran, mengubah interaksi menjadi pendorong adaptasi yang berkelanjutan. Hasilnya adalah lapisan kebijakan dinamis

yang menyeimbangkan strategi yang telah terbukti dengan eksplorasi baru, membantu agen tetap efektif bahkan dalam situasi yang tidak dapat diprediksi. Misalnya, mekanisme penggunaan alat adaptif seperti MeCo menunjukkan bagaimana agen dapat menggunakan sinyal metakognitif untuk memutuskan kapan harus menggunakan alat, yang menggambarkan adaptasi kebijakan dinamis.

Arsitektur

Arsitektur agen yang berevolusi sendiri menawarkan fondasi struktural yang mengintegrasikan model, memori, alat, dan mekanisme umpan balik ke dalam sistem yang terpadu. Karakteristik utama dari arsitektur ini adalah modularitas, yang memungkinkan komponen untuk berevolusi secara independen sambil tetap berkontribusi pada peningkatan sistem secara keseluruhan. Tidak seperti desain statis, arsitektur yang berevolusi sendiri mendukung aliran informasi berkelanjutan dan umpan balik tertutup, sehingga hasil penalaran dan tindakan dapat secara langsung memengaruhi keputusan di masa mendatang. Siklus refleksi, tindakan, dan penyempurnaan ini mengubah arsitektur dari alur kerja tetap menjadi sistem dinamis yang mengoptimalkan diri sendiri.

Evolusi arsitektur dapat terjadi pada tingkat agen tunggal dan multi-agen. Dalam sistem agen tunggal, optimasi sering berfokus pada peningkatan alur kerja dengan menyesuaikan modul internal, menyusun ulang langkah-langkah penalaran, atau bahkan menulis ulang kode untuk mengatasi kelemahan yang diamati. Dalam sistem multi-agen, evolusi arsitektur menekankan koordinasi dan kecerdasan kolektif. Riset menunjukkan bahwa sistem multi-agen dapat mengoptimalkan komposisi tim, strategi komunikasi, dan mekanisme koordinasi, yang mengarah pada perilaku yang muncul seperti alokasi tugas yang efisien atau pemecahan masalah kolaboratif. Yang penting, arsitektur ini bukanlah desain tetap. Arsitektur ini berevolusi sebagai respons terhadap persyaratan tugas, batasan komputasi, dan pengalaman yang terakumulasi. Dengan cara ini, arsitektur bertindak sebagai kerangka dan mesin evolusi diri, memungkinkan agen untuk beradaptasi secara efektif di lingkungan yang semakin kompleks (Gambar 6.1).



Gambar 6.1 Empat modul inti dari agen yang berevolusi sendiri

Memori

Memori adalah pilar fundamental dalam agen yang berevolusi sendiri, karena memungkinkan akumulasi, pengambilan, dan penataan pengalaman masa lalu yang membentuk substrat untuk adaptasi, generalisasi, dan penghindaran kesalahan. Tidak seperti sistem memori konvensional yang terutama melayani penyelesaian tugas jangka pendek, memori agen yang berevolusi sendiri harus mendukung adaptabilitas langsung dan pembelajaran sepanjang hayat. Persyaratan ganda ini membutuhkan arsitektur canggih yang terus menyeimbangkan memori kerja sementara dengan repositori jangka panjang yang persisten. Dalam praktiknya, memori biasanya diorganisasikan menjadi komponen jangka pendek dan jangka panjang. Memori jangka pendek menyediakan landasan kontekstual untuk interaksi agen saat ini, sementara memori jangka panjang menyimpan pengetahuan yang mencakup berbagtugas AI dan episode, termasuk informasi faktual, abstraksi tingkat tinggi, keterampilan prosedural, dan catatan keberhasilan dan kegagalan di masa lalu. Metode modern berbantuan pengambilan memungkinkan agen untuk mengambil memori yang relevan pada saat penalaran, memastikan bahwa pengambilan keputusan didasarkan pada pengalaman yang terakumulasi daripada terbatas pada konteks lokal dari masukan. Secara paralel, teknik seperti pemangkasan, pengelompokan, peringkasan, dan penilaian relevansi membantu mencegah memori jangka panjang menjadi tidak terkendali, sehingga mempertahankan efisiensi selama periode penggunaan yang panjang.

Di luar penyimpanan dan pengambilan statis, agen yang berevolusi sendiri menunjukkan otonomi dalam membentuk isi memori mereka. Mereka memutuskan apa yang akan disimpan, dalam bentuk apa, dan bagaimana mengatur ulang dan memperbarui pengetahuan seiring dengan akumulasi pengalaman baru. Proses reflektif sering digunakan untuk memberi label pada pengalaman, membedakan antara strategi yang berhasil dan tidak berhasil, dan untuk mengabstraksikan keterampilan yang dapat digeneralisasikan atau memadatkan catatan yang sudah usang. Kerangka Kerja Pembelajaran Sepanjang Hayat Berbasis Pengalaman (ELL) menyoroti dinamika ini dengan menekankan restrukturisasi berkelanjutan dari jejak pengalaman menjadi bentuk pengetahuan tingkat tinggi yang dapat digunakan kembali dan diadaptasi dalam skenario baru. Lebih lanjut, survei tersebut menekankan bahwa evolusi memori dapat bersifat internal dan eksternal. Secara internal, agen merestrukturisasi basis pengetahuan mereka melalui konsolidasi reflektif, sementara secara eksternal, mereka dapat berbagi pengalaman dengan agen lain dalam ekosistem multi-agen, sehingga memungkinkan memori kolektif yang mempercepat evolusi kemampuan di seluruh populasi agen.

Alat

Alat eksternal memperluas ruang kemampuan agen yang berevolusi sendiri melampaui apa yang dapat dicapai oleh model internal mereka sendiri. Alat menyediakan akses ke informasi baru, presisi komputasi, aktuasi fisik, atau API khusus yang membuka jalan untuk menyelesaikan tugas yang sebelumnya berada di luar jangkauan. Esensi agen yang berevolusi sendiri tidak hanya terletak pada kemampuan mereka untuk memanggil alat, tetapi juga pada adaptasi dinamis, optimasi, dan perluasan penggunaan alat dari waktu ke waktu.Pada tingkat

paling dasar, agen harus memilih di antara alat yang tersedia, menghasilkan masukan yang sesuai, dan mengintegrasikan keluaran secara mulus ke dalam alur penalaran mereka. Namun, agen yang berevolusi sendiri harus melampaui kompetensi awal ini menuju kemampuan untuk menemukan dan menguasai alat baru. Ini melibatkan pembelajaran untuk membaca dokumentasi, menafsirkan contoh, atau secara aktif bereksperimen dengan panggilan antarmuka hingga pola penggunaan yang andal terbentuk. Ketika tidak ada alat yang sesuai, agen bahkan dapat membangun alat baru atau membungkus fungsionalitas yang ada, sehingga memperluas lingkungan operasional mereka sendiri. Mekanisme tersebut mencerminkan proses akuisisi dan inovasi alat yang berkelanjutan yang sejajar dengan pembelajaran manusia di domain terbuka.

Survei tersebut menekankan bahwa evolusi alat tidak terbatas pada perluasan perangkat alat tetapi juga mencakup penyempurnaan kebijakan penggunaan alat. Agen harus belajar untuk memutuskan dalam kondisi apa alat tersebut harus digunakan, bagaimana menimbang pertimbangan antara biaya, keandalan, dan efisiensi, dan bagaimana memulihkan diri dengan baik ketika alat gagal. Pendekatan terbaru memperkenalkan sinyal metakognitif yang memandu keputusan ini, memastikan bahwa penggunaan alat hanya terjadi ketika meningkatkan kinerja, seperti yang ditunjukkan oleh kerangka kerja MeCo. Seiring waktu, agen dapat membangun hierarki penggunaan alat, menggabungkan beberapa alat, atau mengembangkan strategi cadangan yang memberikan ketahanan dalam lingkungan dinamis. Sama pentingnya adalah integrasi penggunaan alat ke dalam lingkaran umpan balik yang lebih luas yang mendorong evolusi diri. Setiap penggunaan alat menghasilkan hasil yang mungkin berhasil, gagal, atau menghasilkan hasil parsial, dan hasil ini menghasilkan sinyal yang mengalir kembali ke memori, model, dan lapisan kebijakan agen. Melalui siklus ini, agen tidak hanya belajar bagaimana menggunakan alat secara lebih efektif tetapi juga menyempurnakan strategi internal mereka sebagai respons terhadap kinerja alat. Dalam konteks multi-agen, evolusi alat mengambil dimensi tambahan, karena agen dapat berbagi pengetahuan tentang alat, mengamati strategi orang lain, dan secara kolektif memperluas perangkat alat yang mereka akses, sehingga mempercepat lintasan keseluruhan evolusi diri.

6.3 ARSITEKTUR UTAMA AGEN YANG BEREVOLUSI SENDIRI

Transisi dari agen AI statis ke agen yang berevolusi sendiri memerlukan inovasi arsitektur yang memungkinkan peningkatan otonom dan berkelanjutan. Evolusi ini tidak bersifat monolitik tetapi terjadi di seluruh komponen inti agen, modelnya, memorinya, dan alat-alatnya, baik secara individual maupun bersama-sama. Bagian ini menguraikan paradigma arsitektur utama yang memfasilitasi pertumbuhan mandiri ini, mengkategorikannya berdasarkan komponen yang terlibat dalam siklus evolusi.

Evolusi Diri Model

Evolusi diri model menangkap kapasitas mesin penalaran pusat agen, yang paling sering diwujudkan sebagai model bahasa yang besar, untuk memperbaiki parameter, kebijakan pengambilan keputusan, dan keluarannya sendiri berdasarkan pengalaman yang terakumulasi tanpa memerlukan intervensi manusia secara langsung. Ini mewakili dimensi paling mendasar

dari evolusi diri, karena secara langsung menargetkan peningkatan proses kognitif internal agen. Ambisi utamanya adalah untuk memungkinkan model tidak hanya mengeksekusi tugas tetapi juga untuk secara kritis mengevaluasi dan meningkatkan pola penalaran sendiri dari waktu ke waktu, sehingga mengembangkan kapasitas adaptasi otonom yang semakin canggih. Arsitektur yang mendukung evolusi diri model sering kali bergantung pada lingkaran umpan balik metakognitif yang memungkinkan sistem untuk menghasilkan dan bertindak berdasarkan penilaiannya sendiri.

Mekanisme kunci dalam hal ini adalah refleksi diri. Setelah menyelesaikan tugas, model didorong untuk mengevaluasi kinerjanya sendiri, mendeteksi kemungkinan kesalahan dalam penalaran atau perencanaan, dan mengartikulasikan umpan balik eksplisit tentang bagaimana strateginya dapat ditingkatkan. Umpan balik ini tidak dibuang tetapi malah menjadi sumber materi pelatihan, memandu sintesis contoh yang disempurnakan untuk penyempurnaan tambahan atau berfungsi sebagai penyesuaian langsung terhadap strategi penalaran model dalam tugas-tugas mendatang. Siklus pengalaman, refleksi, dan pembelajaran seperti itu menciptakan kondisi untuk pertumbuhan internal iteratif, di mana setiap episode menjadi batu loncatan menuju kinerja yang lebih andal. Kemajuan lebih lanjut dalam lintasan ini adalah pengembangan peningkatan diri generatif. Dalam pendekatan ini, model secara otonom menghasilkan data sintetis, termasuk pasangan masalah-solusi baru atau lintasan penalaran alternatif, yang kemudian dimasukkan ke dalam proses pelatihannya. Dengan menghasilkan materi pembelajarannya sendiri, model dapat memperkuat strategi yang efektif, bereksperimen dengan konsep-konsep baru, dan secara bertahap memperluas keragaman basis pengetahuan internalnya. Strategi ini dapat dikombinasikan dengan algoritma pembelajaran daring yang memungkinkan pembaruan bertahap pada parameter model. Sinyal penguatan, seperti yang diperoleh dari pembelajaran penguatan dengan umpan balik manusia, berfungsi sebagai indikator keberhasilan atau kegagalan yang berharga dan memberikan dasar untuk memandu penyesuaian parameter yang berkelanjutan.

Tantangan arsitektur utama dalam evolusi diri model terletak pada perancangan mekanisme umpan balik yang tetap stabil dan konstruktif dari waktu ke waktu. Meskipun refleksi yang dihasilkan sendiri dan data sintetis memberikan peluang yang kuat untuk pertumbuhan otonom, mereka juga membawa risiko kesalahan yang berlipat ganda dan memperkuat pola penalaran yang salah. Memastikan bahwa siklus evolusi menghasilkan peningkatan berkelanjutan dan bukan pergeseran degeneratif memerlukan penataan kriteria evaluasi yang cermat, integrasi sinyal validasi eksternal, dan desain mekanisme yang menyeimbangkan eksplorasi dengan keandalan. Mengatasi tantangan ini sangat penting untuk mengubah model dari sistem penalaran statis menjadi agen yang mampu melakukan perbaikan diri tanpa batas.

Evolusi Diri Bersama Model-Memori

Arsitektur evolusi diri model-memori bersama menekankan bahwa model dan sistem memorinya tidak dapat berkembang secara terisolasi tetapi harus maju bersama untuk mendukung pembelajaran yang berkelanjutan dan kuat. Model membentuk isi dan struktur dari apa yang disimpan dalam memori, sementara memori yang berevolusi, pada gilirannya,

menyediakan fondasi kontekstual yang memperkaya penalaran dan peningkatan diri model di masa depan. Ketergantungan ini menciptakan dinamika dua arah di mana setiap komponen berfungsi sebagai pendorong dan penerima manfaat dari evolusi komponen lainnya. Di pusat proses ini terletak desain mekanisme memori yang dinamis dan dikurasi secara selektif. Alih-alih menyimpan pengalaman mentah tanpa pandang bulu, model melakukan langkah reflektif setelah setiap tugas untuk mengevaluasi hasil dan mengekstrak pelajaran penting. Pelajaran ini sering kali mencakup strategi kunci yang terbukti berhasil, kesalahan yang menghambat kinerja, dan wawasan yang lebih luas yang mungkin relevan untuk pemecahan masalah di masa mendatang.

Pengetahuan yang diekstrak kemudian diubah menjadi entri terstruktur, dianotasi dengan metadata seperti kategori tugas atau indikator kinerja, dan disimpan dalam memori jangka panjang. Dengan cara ini, memori berevolusi dari arsip pasif interaksi masa lalu menjadi repositori aktif dan terstruktur dari pengalaman yang disaring, menyerupai grafik pengetahuan yang mengkodekan kumpulan kebijaksanaan yang dipelajari agen yang terus berkembang. Namun, memori bukan hanya mekanisme penyimpanan tetapi juga sistem pengambilan yang canggih yang secara aktif berkontribusi pada penalaran yang sedang berlangsung. Ketika dihadapkan dengan tantangan baru, model tersebut menggunakan memori untuk mencari tidak hanya kecocokan yang tepat tetapi juga pengalaman analogis yang dapat menawarkan strategi yang dapat ditransfer. Melalui proses ini, model tersebut mengadaptasi pengetahuan masa lalu ke konteks baru, mencapai tingkat generalisasi yang melampaui hafalan. Kemajuan terbaru dalam generasi yang ditingkatkan dengan pengambilan dan teknik pencarian vektor padat telah secara signifikan memperkuat kapasitas ini dengan memungkinkan kueri memori yang efisien dan kaya konteks. Pertukaran berkelanjutan antara model dan memori ini membentuk siklus yang baik: model yang lebih baik menghasilkan memori yang lebih terstruktur dan berharga, sementara repositori memori yang diperkaya memberdayakan model untuk bernalar dengan kedalaman, fleksibilitas, dan pandangan ke depan yang lebih besar.

Evolusi Diri Bersama Model-Alat

Prinsip evolusi diri juga dapat diperluas ke interaksi antara model dan alat, di mana alat digunakan bukan hanya sebagai sumber daya eksternal untuk eksekusi tugas tetapi sebagai instrumen aktif yang meningkatkan pembelajaran dan adaptasi model. Dalam arsitektur ini, alat berfungsi sebagai perluasan jangkauan operasional agen dan sebagai katalis untuk mengumpulkan data dan umpan balik baru yang memandu penyempurnaan model itu sendiri. Hasilnya adalah sistem di mana penggunaan alat yang strategis mendorong pertumbuhan kognitif dan perluasan operasional. Salah satu kemampuan utama dari paradigma ini adalah kapasitas untuk penemuan dan pembelajaran alat secara otomatis. Agen dapat memanfaatkan modelnya untuk mempelajari dokumentasi alat, menafsirkan spesifikasi API, atau bereksperimen langsung dengan fungsionalitas baru. Melalui proses tersebut, agen secara bertahap mengintegrasikan alat-alat baru ke dalam repertoarnya, sehingga memperluas jangkauan tindakan yang dapat dilakukannya tanpa pengawasan manusia. Akuisisi alat yang

dinamis ini memperluas ruang tindakan model dan berkontribusi langsung pada evolusi adaptifnya.

Selain penemuan, alat juga memainkan peran yang sangat penting dalam akuisisi data dan pembuatan umpan balik. Misalnya, akses ke sistem pencarian web memungkinkan agen untuk mengambil informasi terbaru dan dengan demikian mengatasi keterbatasan pengetahuan yang diwarisi dari pelatihan awal. Demikian pula, penggunaan alat komputasi seperti penerjemah kode memberikan peluang bagi agen untuk menguji hipotesis, memverifikasi prediksi, dan menerima umpan balik yang tepat berdasarkan hasil yang dapat dieksekusi daripada evaluasi linguistik saja. Bentuk umpan balik eksternal ini sangat berharga, karena diintegrasikan ke dalam lingkaran refleksi diri agen dan digunakan untuk menyempurnakan strategi penalaran, kebijakan, dan struktur pengetahuan internalnya.

Merancang arsitektur yang memungkinkan ko-evolusi model-alat yang efektif membutuhkan pengelolaan yang cermat terhadap interaksi antara keluaran alat dan adaptasi model. Umpan balik yang diperoleh dari alat harus diuraikan secara akurat, dikontekstualisasikan, dan diterjemahkan menjadi sinyal yang berkontribusi secara bermakna pada pembelajaran. Ketika dieksekusi secara efektif, integrasi ini membangun lingkungan yang kaya akan umpan balik di mana alat tidak hanya memperluas kemampuan operasional agen tetapi juga mendorong siklus iteratif peningkatan diri. Dengan demikian, evolusi diri model-alat bersama memberikan fondasi bagi agen yang dapat terus beradaptasi dengan lingkungan baru, memperluas repertoar pemecahan masalah mereka, dan berevolusi menjadi sistem yang lebih serbaguna dan otonom.

Evolusi Diri Bersama Model-Memori-Alat

Tahap paling canggih dari desain agen yang berevolusi sendiri menyatukan model, memori, dan alat ke dalam arsitektur terpadu yang berfungsi sebagai sistem pembelajaran yang sepenuhnya otonom dan adaptif. Dalam kerangka kerja terintegrasi seperti itu, kapasitas penalaran model, pengetahuan yang terakumulasi yang tersimpan dalam memori, dan kemampuan interaksi yang disediakan oleh alat tidak berkembang secara independen tetapi berevolusi bersama secara erat dan saling memperkuat. Sinergi ini mengubah agen dari pelaksana tugas spesifik menjadi entitas dinamis yang mampu melakukan peningkatan diri secara berkelanjutan dalam jangka waktu yang panjang. Dalam arsitektur ini, setiap aspek riwayat interaksi agen dimanfaatkan untuk pembelajaran reflektif. Model memproses rencana, pemanggilan alat, keluaran alat, dan hasil tugas dalam analisis retrospektif yang komprehensif. Analisis ini tidak hanya menilai apakah hasil akhir berhasil tetapi juga mengevaluasi seberapa efektif alat dipilih dan diterapkan, bagaimana keluaran alat diinterpretasikan, dan bagaimana pengambilan memori memengaruhi jalur penalaran. Wawasan yang diperoleh dari proses ini diubah menjadi memori terstruktur yang tidak hanya menangkap hasil faktual tetapi juga meta-pengetahuan, seperti keandalan relatif alat, strategi yang menghasilkan pengambilan memori yang efektif, dan alur kerja yang menghasilkan hasil yang sukses.

Melalui mekanisme ini, agen mengembangkan kapasitas pembelajaran tingkat tinggi. Seiring waktu, ia dapat menyimpulkan, misalnya, bahwa alat tertentu secara konsisten mengungguli alternatif untuk kelas masalah tertentu dan mengkodekan preferensi ini ke dalam

memori untuk pengambilan keputusan di masa mendatang. Demikian pula, sistem dapat mengidentifikasi jenis kueri memori mana yang paling efektif untuk tugas penalaran yang berbeda, membedakan antara pengambilan berbasis kata kunci dan pendekatan berbasis analogi. Wawasan ini terakumulasi menjadi pengetahuan prosedural yang memungkinkan agen untuk secara adaptif menyempurnakan strateginya sendiri. Di luar pengoptimalan alur kerja yang ada, arsitektur terintegrasi juga mendukung sintesis kreatif alur kerja baru, di mana model belajar untuk menggabungkan alat dengan pengetahuan yang diambil dalam konfigurasi baru yang terbukti efektif untuk tantangan yang muncul.

Memori yang berkembang dalam konteks ini menjadi jauh lebih dari sekadar repositori hasil tugas. Ia berfungsi sebagai basis data yang terus berkembang dari pengetahuan khusus domain, taktik prosedural, dan wawasan metakognitif ke dalam pola operasional agen itu sendiri. Memori berlapis ini memungkinkan agen tidak hanya mengingat apa yang telah dipelajari tetapi juga memahami bagaimana ia belajar secara paling efektif, menciptakan lingkaran peningkatan rekursif yang memperkuat kompetensi dan adaptabilitas. Dengan mengaitkan erat evolusi model, memori, dan alat, arsitektur ini membangun sistem lingkaran tertutup di mana setiap komponen memperkuat kemampuan komponen lainnya. Hasilnya adalah agen yang mampu tumbuh secara berkelanjutan, beradaptasi secara otonom terhadap lingkungan baru, dan tangguh dalam menghadapi kompleksitas dunia terbuka. Integrasi semacam itu mewakili puncak paradigma evolusi diri, memberikan fondasi bagi agen yang tidak hanya reaktif tetapi benar-benar mengarahkan diri sendiri dalam upaya peningkatan berkelanjutan.

6.4 STRATEGI UTAMA AGEN YANG BEREVOLUSI SENDIRI

Dimensi penting dari agen yang berevolusi diri adalah serangkaian strategi pembelajaran yang mengubah interaksi dan pengamatan mentah menjadi peningkatan berkelanjutan dalam kebijakan, memori, alat, dan alur kerja. Strategi ini menentukan baik jenis sinyal maupun proses agen, seperti kritik tekstual, imbalan skalar, atau hasil lingkungan, dan metode yang digunakan untuk memanfaatkan sinyal-sinyal ini, termasuk imitasi, penguatan, distilasi, dan pencarian populasi. Bahan bacaan mengidentifikasi dua keluarga strategi utama yang saling tegak lurus: pendekatan berbasis imitasi, yang bergantung pada demonstrasi berkualitas tinggi sebagai sinyal pelatihan utama, dan pendekatan berbasis interaksi, yang memperoleh sinyal pembelajaran dari keterlibatan agen dengan lingkungannya.

Evolusi Diri Berbasis Imitasi

Evolusi diri berbasis imitasi memperlakukan eksekusi yang berhasil, baik yang dihasilkan oleh pakar manusia, model "guru" yang lebih kuat, atau upaya terbaik agen di masa lalu, sebagai sinyal pembelajaran kanonik. Paradigma ini sangat ampuh untuk meningkatkan kompetensi di domain di mana eksplorasi tanpa panduan tidak efisien atau berbahaya, dan menyediakan jalur yang efisien dalam hal sampel untuk mempelajari strategi multi-langkah yang sulit ditemukan. Bahan bacaan mengelompokkan beberapa modalitas operasional di bawah payung ini: kloning perilaku, distilasi, imitasi diri, berbagi demonstrasi lintas agen, dan alur kerja demonstrasi hibrida. Kloning perilaku (atau imitasi terawasi) membingkai lintasan

yang berhasil sebagai contoh terawasi: pasangan keadaan/tindakan (atau perintah/keluaran) dikumpulkan ke dalam kumpulan data yang digunakan untuk menyempurnakan kebijakan agen atau menghasilkan templat perintah. Pendekatan ini mudah dan efektif ketika demonstrasi berlimpah dan representatif, dan dapat diimplementasikan baik melalui penyempurnaan terawasi parameter penuh atau adaptasi efisien parameter seperti LoRA untuk model besar. Kloning perilaku telah digunakan sebagai dasar dalam banyak alur kerja agen di mana solusi berkualitas tinggi tersedia atau dapat disintesis.

Distilasi pengetahuan meningkatkan imitasi dengan memungkinkan agen untuk menanyakan model "guru" yang lebih mampu tetapi mahal untuk label, penjelasan, atau jejak penalaran, dan kemudian mengompres kemampuan tersebut ke dalam model siswa yang lebih murah. Distilasi adalah jalur pragmatis untuk peningkatan luring berkelanjutan: ia menghasilkan model yang mendekati kompetensi guru sambil mempertahankan jejak waktu eksekusi yang dapat dikelola. Dalam konteks agen, distilasi dapat beroperasi pada berbagai tingkat representasi, jawaban akhir, rantai pemikiran, dan bahkan kebijakan pemilihan alat, memungkinkan agen siswa untuk menginternalisasi heuristik pengambilan keputusan yang lebih kaya daripada imitasi tindakan sederhana. Pembelajaran imitasi diri memainkan peran komplementer: dengan menyusun "aula kemasyhuran" dari episode terbaik agen itu sendiri, sistem berulang kali menggunakan kembali lintasan yang paling sukses sebagai data pelatihan berkualitas tinggi. Imitasi diri mengurangi kelupaan yang dahsyat terhadap perilaku yang langka namun berharga dan menstabilkan pembelajaran dengan memperkuat strategi yang telah terbukti berhasil secara empiris. Metode seperti STaR dan alur kerja pelatihan mandiri berturut-turut menggambarkan bagaimana agen dapat secara iteratif meningkatkan kemampuan penalaran dan penggunaan alat dari keluarannya sendiri yang telah divalidasi.

Pembelajaran demonstrasi lintas agen dan skema hibrida memperluas paradigma imitasi di luar peningkatan mandiri agen tunggal. Dalam ekosistem multi-agen atau multi-model, agen berbagi episode, verifikator, dan demonstrasi yang sukses untuk mempercepat pembelajaran kolektif; alur kerja hibrida menggabungkan contoh yang dihasilkan sendiri dengan data ahli yang bersumber dari luar dan mekanisme seleksi yang dipandu kepercayaan untuk memaksimalkan kekokohan dan keragaman himpunan pelatihan. Mekanisme ini sangat berharga di ranah di mana demonstrasi berkualitas tinggi tidak terdistribusi secara merata di berbagtugas AI atau di mana agen yang berbeda mengkhususkan diri dalam keterampilan yang berbeda. Terlepas dari kekuatannya, evolusi berbasis imitasi bukanlah obat mujarab: efektivitasnya sangat bergantung pada kualitas, cakupan, dan keragaman lintasan contoh. Ketika demonstrasi ahli langka, bising, atau bias, imitasi dapat terlalu sesuai dengan pola yang suboptimal; oleh karena itu, riset terbaru menekankan seleksi yang dipandu oleh verifikator, penyaringan berbasis kepercayaan, dan strategi pengambilan sampel adaptif untuk membuat imitasi kuat dalam pengaturan yang realistis.

Evolusi Diri Berbasis Interaksi

Evolusi diri berbasis interaksi mencakup strategi di mana agen memperoleh pengetahuan dan menyempurnakan perilaku mereka melalui keterlibatan dengan lingkungan mereka, khususnya dalam pengaturan di mana demonstrasi eksplisit atau data berlabel tidak

tersedia. Dalam paradigma ini, agen memperlakukan dunia eksternal atau model internalnya sebagai sumber umpan balik, menggunakan konsekuensi tindakannya untuk memandu pembelajaran dan penyempurnaan kebijakan. Sinyal umpan balik dapat mengambil berbagai bentuk, termasuk anotasi tekstual, imbalan skalar, atau indikator yang lebih implisit yang berasal dari hasil tugas atau interaksi manusia. Bahan bacaan umumnya mengklasifikasikan metode ini menurut asal dan sifat umpan balik, mengidentifikasi tiga kategori utama: motivasi intrinsik, di mana agen menghasilkan sinyal imbalan internal untuk mendorong eksplorasi; tujuan lingkungan, di mana imbalan diberikan secara eksternal oleh lingkungan; dan imbalan implisit, yang tertanam dalam dinamika tugas atau dalam respons kolaborator manusia.

Interaksi Imbalan Internal

Mekanisme imbalan internal mewakili pendekatan inti terhadap motivasi intrinsik, yang memberikan agen tujuan yang dihasilkan sendiri yang memandu eksplorasi dan perolehan keterampilan secara independen dari insentif spesifik tugas eksternal. Pembelajaran yang didorong oleh rasa ingin tahu merupakan contoh pendekatan ini, karena agen menerima imbalan intrinsik karena menemukan keadaan baru atau karena mengurangi kesalahan prediksi dalam model dunia internal. Mekanisme ini mendorong eksplorasi sistematis dan penemuan kemampuan progresif yang dapat ditransfer secara luas di berbagtugas AI. Kerangka kerja penemuan keterampilan tanpa pengawasan lebih lanjut memperluas paradigma ini dengan memberi imbalan pada keragaman perilaku, yang mengarah pada munculnya primitif yang dapat digunakan kembali yang kemudian dapat disusun untuk mengatasi tujuan hilir.

Lingkaran imbalan yang dihasilkan secara internal ini sangat berharga karena mengurangi ketergantungan pada imbalan eksternal yang dirancang dengan cermat, memungkinkan agen untuk melatih dan menyempurnakan kompetensi di lingkungan terbuka dan tidak terstruktur. Studi terbaru menggarisbawahi kegunaan metrik kepercayaan internal, evaluasi diri berbasis ansambel, dan siklus pelatihan yang memberi imbalan diri sebagai sumber supervisi yang terukur dan kuat, menunjukkan bahwa agen dapat mencapai peningkatan otonom yang bermakna bahkan tanpa panduan eksplisit. Dengan memanfaatkan evolusi diri berbasis interaksi, agen mampu mengembangkan perilaku yang semakin canggih yang didasarkan pada keterlibatan pengalaman mereka sendiri dengan lingkungan. Paradigma ini tidak hanya memfasilitasi perolehan keterampilan yang relevan dengan tugas tetapi juga mendorong generalisasi, adaptabilitas, dan munculnya kemampuan penalaran tingkat tinggi, memberikan dasar bagi agen yang dapat terus meningkatkan diri di domain yang kompleks dan dinamis.

Interaksi Imbalan Eksternal

Imbalan eksternal berasal dari sinyal objektif yang diberikan oleh lingkungan atau oleh evaluator tugas, seperti label keberhasilan dan kegagalan, skor kinerja numerik, atau respons dari alat eksternal. Sinyal-sinyal ini memberikan kriteria yang jelas dan terukur untuk memandu perilaku agen, memungkinkan optimasi langsung menuju tujuan yang terdefinisi dengan baik, termasuk navigasi web yang sukses, eksekusi program yang akurat, atau penyelesaian tugas di lingkungan interaktif. Pembelajaran penguatan (RL) dan penyempurnaan

berbasis imbalan (penyempurnaan terkondisi imbalan) merupakan kerangka kerja metodologis utama untuk memanfaatkan imbalan eksternal pada agen yang berevolusi sendiri. Namun, dalam praktiknya, metode-metode ini harus menghadapi masalah seperti umpan balik yang jarang, penugasan kredit yang tertunda, dan ketidakstabilan pelatihan dengan model berkapasitas tinggi. Akibatnya, metode-metode ini sering kali ditambah dengan inovasi prosedural seperti struktur imbalan hierarkis, model evaluasi berbasis proses, atau desain kurikulum yang secara bertahap mengekspos agen pada tugas-tugas yang semakin kompleks. Misalnya, formulasi hierarkis yang menggabungkan sinyal hasil tingkat episode dengan imbalan pembentukan tingkat langkah telah terbukti meningkatkan penugasan kredit dalam pengambilan keputusan jangka panjang.

Demikian pula, mekanisme verifikasi hasil, yang menggabungkan evaluasi eksternal ke dalam siklus pelatihan, membantu memastikan bahwa kebijakan yang dipelajari selaras dengan tujuan yang ditentukan manusia dan mempertahankan keandalan di bawah pergeseran distribusi. Oleh karena itu, banyak sistem agen kontemporer mengadopsi strategi hibrida yang mengintegrasikan RL dengan penyempurnaan terawasi dan verifikasi eksternal, memungkinkan agen untuk menyeimbangkan stabilitas imitasi dengan kemampuan adaptasi penguatan. Dengan mendasarkan evolusi pada sinyal eksternal yang eksplisit, agen mampu menyelaraskan adaptasi mereka dengan tujuan kinerja yang terukur dan dapat diinterpretasikan, sehingga memfasilitasi peningkatan yang dapat diprediksi dan terarah pada tujuan. Kerangka kerja ini sangat efektif di domain di mana metrik keberhasilan yang jelas ada, seperti sintesis program, kontrol robot, atau simulasi multi-agen skala besar, dan tetap menjadi salah satu paradigma yang paling banyak digunakan untuk menyusun evolusi diri agen.

Interaksi Imbalan Implisit

Dalam banyak skenario dunia nyata, imbalan numerik eksplisit tidak tersedia atau sangat mahal untuk diperoleh, dan agen harus menyimpulkan sinyal pembelajaran dari pengamatan tidak langsung, hasil tugas, atau pola interaksi manusia. Imbalan implisit mengubah hasil yang dapat diamati tersebut menjadi sinyal pelatihan yang informatif, memungkinkan adaptasi otonom bahkan tanpa tujuan yang didefinisikan secara formal. Misalnya, refleksi diri berbasis hasil mengubah keberhasilan dan kegagalan tugas menjadi sinyal imbalan implisit dengan mengevaluasi secara retrospektif apakah tindakan berkontribusi pada pencapaian tujuan dan dengan mengekstrak pola kesalahan atau strategi korektif, yang kemudian dapat dimasukkan ke dalam memori agen dan data pelatihan. Umpan balik manusia memberikan sumber pengawasan implisit yang kaya lainnya, mulai dari koreksi eksplisit hingga isyarat yang lebih halus seperti tingkat keterlibatan, kepuasan, atau preferensi yang diungkapkan. Sinyal-sinyal ini dapat dimodelkan dan disaring menjadi fungsi imbalan yang dipelajari yang membentuk keselarasan agen, secara efektif memungkinkan agen untuk belajar dari interaksi sosial yang bernuansa tanpa bergantung pada imbalan spesifik tugas yang berskala besar.

Riset terbaru juga menunjukkan bahwa model bahasa besar dapat memanfaatkan sinyal internal atau endogen, termasuk isyarat skalar yang tertanam dalam konteks, kemungkinan tingkat unit, atau perkiraan ketidakpastian, sebagai imbalan semu. Mekanisme

tersebut memungkinkan model untuk memulai adaptasi menggunakan struktur prediktifnya sendiri, menciptakan sinyal penguatan internal bahkan ketika tidak ada umpan balik eksternal eksplisit yang tersedia. Perspektif ini memperluas cakupan evolusi agen, memungkinkan adaptasi dalam domain terbuka dan kurang spesifik di mana rekayasa imbalan formal tidak praktis. Meskipun demikian, strategi imbalan implisit bukannya tanpa tantangan. Strategi ini membutuhkan mekanisme inferensi yang canggih untuk secara andal mengubah pengamatan tidak langsung menjadi target pembelajaran yang dapat ditindaklanjuti, dan harus dengan hati-hati menyeimbangkan eksplorasi dengan penyesuaian berlebihan terhadap sinyal palsu. Terlepas dari kesulitan ini, interaksi imbalan implisit merupakan pelengkap yang ampuh untuk penguatan eksternal eksplisit. Dengan memanfaatkan isyarat lingkungan dan interaksi yang bernuansa, agen memperoleh kapasitas untuk secara fleksibel meningkatkan diri di domain di mana keberhasilan tidak dapat direduksi menjadi skor numerik sederhana. Dalam hal ini, imbalan implisit memperluas batas sistem yang berevolusi sendiri, melengkapi mereka untuk beroperasi di lingkungan yang kompleks, dinamis, dan kurang terstruktur.

Strategi berbasis imitasi dan interaksi saling melengkapi dan bukan saling eksklusif. Dalam praktiknya, agen yang berevolusi sendiri dengan kinerja tinggi menggabungkan keduanya: imitasi menyediakan inisialisasi yang efisien dalam pengambilan sampel dan heuristik yang stabil, sementara pembelajaran berbasis interaksi memungkinkan adaptasi berkelanjutan dan penemuan kemampuan baru. Pilihan desain tentang apa yang akan ditiru, dorongan internal mana yang akan diwujudkan, bagaimana membangun model imbalan eksternal, dan bagaimana menyimpulkan sinyal implisit mengatur efisiensi sampel, keamanan, dan generalisasi agen. Praktik terbaik yang muncul menekankan alur kerja hibrida, mekanisme seleksi yang sadar akan kepercayaan diri, dan arsitektur modular yang memungkinkan sinyal evolusi yang berbeda untuk memperbarui komponen yang sesuai dengan pengaruh yang terkontrol.

6.5 STUDI KASUS: AGEN YANG BEREVOLUSI SENDIRI

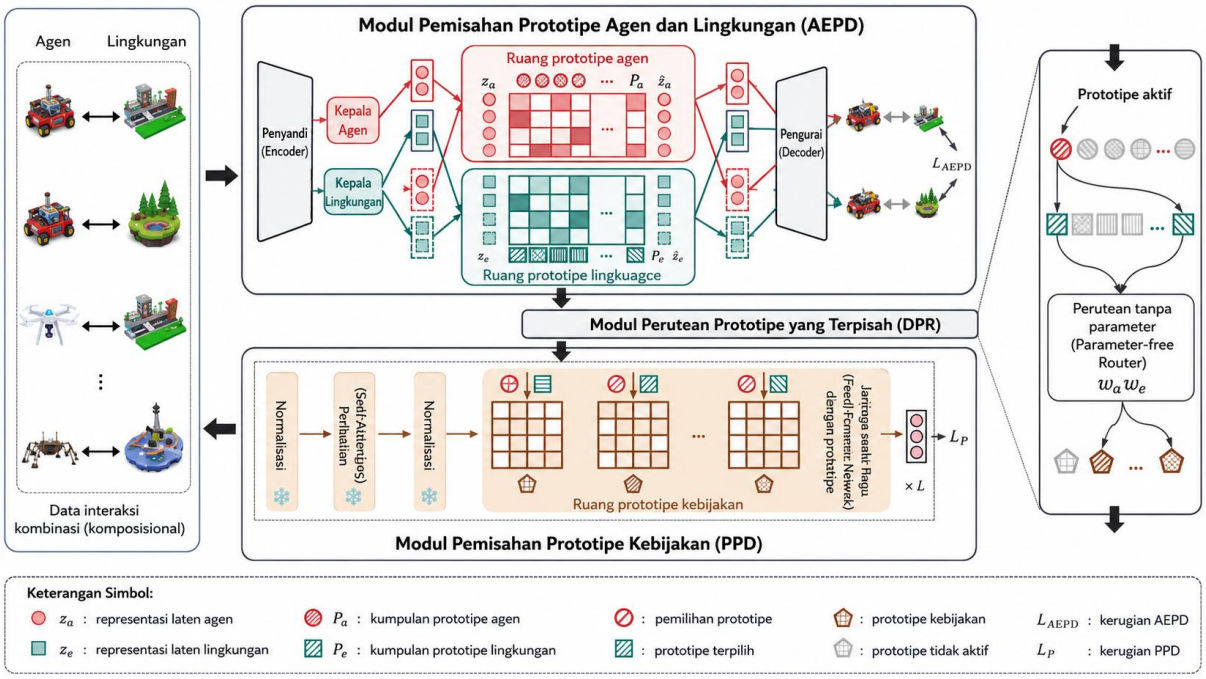
Sebagai contoh nyata dari sistem berwujud yang berevolusi sendiri, kerangka kerja Campuran Prototipe Terurai (MoDP) mengilustrasikan bagaimana representasi modular, adaptasi arsitektur yang ditargetkan, dan penyempurnaan prototipe daring dapat bersama-sama menghasilkan agen yang terus-menerus mengatur ulang ruang perilakunya sebagai respons terhadap komposisi perwujudan dan lingkungan yang baru. Tidak seperti pembelajaran penguatan ujung-ke-ujung konvensional atau kerangka kerja berbasis imitasi, yang seringkali mengaitkan faktor interaksi heterogen ke dalam parameterisasi monolitik, MoDP secara eksplisit memisahkan struktur masalah dan memperkenalkan modularitas pada tingkat representasi dan arsitektur kebijakan. Pemisahan ini memungkinkan sistem untuk memperoleh blok bangunan yang dapat digunakan kembali yang dapat disempurnakan dan dikombinasikan secara bertahap untuk mengakomodasi komposisi agen dan lingkungan yang baru. Dengan cara ini, MoDP tidak hanya meningkatkan generalisasi komposisi tetapi juga meletakkan dasar untuk adaptasi berkelanjutan, yang merupakan ciri khas evolusi diri.

Desain Kerangka Kerja

Kerangka kerja ini beroperasi dengan menguraikan alur observasi-ke-aksi menjadi tiga modul yang berbeda namun saling terkait, yang masing-masing memainkan peran komplementer dalam memisahkan, mengkhususkan, dan menyusun kembali faktor-faktor interaksi. Modul pemisahan prototipe agen dan lingkungan (AEPD) menerima masukan berupa penyematan lintasan dari pasangan agen dan lingkungan tertentu, dan memetakannya ke dalam dua ruang laten terpisah, satu menangkap dinamika spesifik agen dan yang lainnya mengkodekan kendala spesifik lingkungan. Dengan memproyeksikan vektor laten ini ke himpunan prototipe kompak melalui kombinasi cembung berbobot kesamaan, modul ini memastikan bahwa faktor interaksi direpresentasikan oleh repertoar primitif yang terbatas dan dapat digunakan kembali. Pemisahan ini memberikan bias struktural yang diperlukan untuk generalisasi komposisional, karena prototipe agen yang sama dapat dikombinasikan kembali dengan prototipe lingkungan dari konteks yang belum pernah dilihat sebelumnya untuk menghasilkan representasi yang masuk akal untuk konfigurasi baru. Berdasarkan representasi yang terpisah ini, modul pemisahan prototipe kebijakan (PPD) mengekstrak pustaka sub-jaringan fungsional dari jaringan utama yang telah dilatih sebelumnya menggunakan pemangkasan sensitif koneksi.

Setiap sub-jaringan, atau prototipe kebijakan, dihilangkan dari jaringan utama sesuai dengan sensitivitas parameternya terhadap perilaku tertentu, sehingga mempertahankan komponen fungsional yang paling relevan sambil menghilangkan redundansi. Proses ini mengubah ruang kebijakan dari parameterisasi monolitik menjadi kumpulan sub-kebijakan yang dapat diinterpretasikan secara modular, yang masing-masing mengkhususkan diri dalam himpunan bagian perilaku yang selaras dengan prototipe yang berbeda. Tahap terakhir, perutean prototipe dinamis (DPR), memperkenalkan mekanisme ringan untuk komposisi kebijakan. Dengan diberikan aktivasi prototipe dari modul AEPD, DPR menghitung penugasan lunak atas prototipe kebijakan yang tersedia dan merute masukan melalui campuran sub-jaringan yang sesuai, secara efektif menyusun kebijakan komposit tanpa memerlukan pelatihan ulang global.

Dengan cara ini, mekanisme perutean memungkinkan adaptasi cepat terhadap pasangan agen dan lingkungan yang belum pernah dilihat sebelumnya dengan melakukan interpolasi antara sub-kebijakan khusus secara bebas parameter. Secara keseluruhan, ketiga modul ini mengoperasionalkan prinsip evolusi diri modular: pemisahan faktor heterogen, spesialisasi komponen fungsional, dan rekomposisi fleksibel dari primitif yang dipelajari. Desain ini membatasi ruang interaksi yang sebelumnya sulit diatasi menjadi varietas laten terstruktur dengan mekanisme eksplisit untuk adaptasi inkremental, memberikan stabilitas dan plastisitas dalam pembelajaran jangka panjang dan dengan demikian mendukung evolusi diri berkelanjutan dari agen terwujud (Gambar 6.2).



Gambar 6.2 Kerangka kerja MoDP yang diusulkan

Mekanisme Pembelajaran

Kekuatan penting MoDP terletak pada aturan pembaruan lokal dan modularnya, yang memungkinkan penyempurnaan berkelanjutan tanpa gangguan yang merusak atau beban komputasi dari pelatihan ulang global. Alih-alih memerlukan pembaruan parameter di seluruh model, MoDP memperkenalkan mekanisme terstruktur untuk adaptasi yang ditargetkan, sehingga mencapai keseimbangan antara stabilitas dan plastisitas. Prototipe yang terkait dengan agen dan lingkungan dioptimalkan melalui tujuan rekonstruksi komposisional, diperkuat dengan istilah regularisasi yang mendorong representativitas, keunikan, dan kelangkaan. Formulasi ini memastikan bahwa prototipe tetap beragam dan kompak, mencegah redundansi sambil tetap mempertahankan cakupan yang cukup dari ruang interaksi. Dengan mempertahankan keragaman ini, kerangka kerja memungkinkan komposisi baru agen dan lingkungan untuk direpresentasikan secara efektif tanpa memerlukan reorganisasi penuh dari struktur yang dipelajari. Prototipe kebijakan mengikuti mekanisme pembaruan yang komplementer tetapi berbeda. Mereka didefinisikan oleh sensitivitas koneksi, sebuah prinsip yang didasarkan pada bahan bacaan pemangkasan dan kompresi jaringan, yang mengidentifikasi sub-jaringan dari model tulang punggung sesuai dengan kepentingannya untuk perilaku spesifik. Sub-jaringan ini dipertahankan melalui masker biner, yang berevolusi seiring akumulasi pengalaman tambahan.

Dengan cara ini, MoDP secara efektif memperlakukan setiap prototipe kebijakan sebagai sub-jaringan fungsional yang dapat disempurnakan dari waktu ke waktu, memungkinkan munculnya pakar khusus yang selaras dengan pola interaksi yang berulang. Yang penting, proses ini tidak statis; masker disempurnakan secara dinamis seiring perubahan distribusi tugas dan lingkungan, memastikan adaptasi berkelanjutan tanpa membuang kompetensi yang telah diperoleh sebelumnya. Semua prototipe dan masker terkaitnya

diperbarui lebih lanjut dengan skema agregasi berbasis momentum, yang menimbang pengaruh pengalaman baru dengan aktivasi prototipe. Mekanisme ini memungkinkan sistem untuk menyeimbangkan sensitivitas jangka pendek terhadap sinyal baru dengan stabilitas jangka panjang, sehingga mendukung akumulasi pengetahuan inkremental tanpa osilasi atau kelupaan katastrofik. Secara keseluruhan, mekanisme-mekanisme ini menggambarkan prinsip utama dari sistem yang berevolusi sendiri: kemampuan untuk mengakumulasi elemen-elemen dasar yang terspesialisasi, memperhalusnya secara selektif, dan menyesuaikan komposisinya dengan keadaan baru dengan cara yang tangguh dan fleksibel.

Evaluasi

Evaluasi empiris MoDP menggarisbawahi manfaat praktis dari mekanisme pembelajaran ini. Di seluruh tolok ukur lintas perwujudan standar, MoDP secara konsisten melampaui dasar transformator kuat HPT, mencapai kerugian validasi yang lebih rendah dan pengurangan besaran gradien pada tahap pelatihan selanjutnya. Pengurangan besaran gradien memberikan bukti bahwa penyelarasan prototipe mengurangi interferensi berbahaya antar tugas, tantangan yang terus-menerus dalam pembelajaran berkelanjutan dan transfer. Ini menunjukkan bahwa prototipe yang terpisah tidak hanya meningkatkan efisiensi tetapi juga mengurangi ketidakstabilan yang biasanya menyertai adaptasi lintas domain. Lebih penting lagi, di bawah rezim generalisasi komposisional yang sengaja dibangun yang melibatkan 25 pasangan agen dan lingkungan, MoDP menunjukkan peningkatan substansial di semua tingkat kesulitan generalisasi. Kerangka kerja ini mencapai tingkat keberhasilan yang jauh lebih tinggi dalam kasus yang konsisten dengan domain dan menunjukkan ketahanan yang unggul ketika ditugaskan untuk melakukan generalisasi ke agen yang belum pernah dilihat sebelumnya, lingkungan yang belum pernah dilihat sebelumnya, atau komposisi agen dan lingkungan yang sepenuhnya baru.

Peningkatan yang diamati cukup signifikan: dari 0,59 menjadi 0,73 dalam domain yang sama, dari 0,36 menjadi 0,55 dalam kasus lintas agen, dari 0,39 menjadi 0,54 dalam kasus lintas lingkungan, dan dari 0,22 menjadi 0,38 dalam skenario lintas domain yang paling menantang. Hasil ini memberikan bukti yang meyakinkan bahwa prototipe yang terpisah dan spesialisasi sub-jaringan modular memungkinkan bentuk transfer pengetahuan yang lebih terukur, mengubah kebaruan komposisional dari sumber gangguan menjadi peluang untuk penggunaan kembali yang terstruktur. Di luar metrik kinerja mentah, studi ablasi memberikan bukti penting untuk memisahkan kontribusi pilihan desain individu dalam MoDP dan dengan demikian memvalidasi sifat evolusi dirinya. Salah satu pengamatan yang paling mencolok adalah peran kerugian rekonstruksi komposisional. Ketika tujuan ini dihilangkan, kinerja turun secara signifikan di seluruh tolok ukur generalisasi komposisional, menegaskan bahwa pemodelan dan penegakan pemisahan secara eksplisit pada tingkat representasi sangat diperlukan. Hasil ini menggarisbawahi pentingnya membatasi ruang representasi dengan cara yang menangkap faktor-faktor spesifik agen dan lingkungan secara independen, daripada membiarkan mereka runtuh menjadi penyematan yang saling terkait yang mengganggu transfer.

Demikian pula, penghilangan mekanisme prototipe kebijakan sangat melemahkan kinerja, menunjukkan bahwa peningkatan tersebut tidak hanya berasal dari peningkatan kapasitas parameter atau redundansi insidental. Sebaliknya, peningkatan tersebut berasal dari penataan proses adaptasi yang berprinsip, di mana sub-jaringan modular secara eksplisit diselaraskan dengan pola interaksi yang berulang. Dengan mengaitkan prototipe ke sub-jaringan yang diidentifikasi melalui sensitivitas koneksi, MoDP memastikan bahwa setiap pakar kebijakan khusus mempertahankan koherensi fungsional, memungkinkan mereka untuk digunakan kembali atau disempurnakan secara selektif seiring evolusi distribusi tugas. Mekanisme ini mencerminkan temuan dalam riset pembelajaran berkelanjutan, di mana modularisasi dan sparsitas sering mengurangi interferensi dan memungkinkan transfer yang lebih stabil.

Analisis kardinalitas prototipe lebih lanjut menjelaskan keseimbangan yang rumit antara efisiensi dan ekspresivitas. Hasil menunjukkan bahwa mempertahankan kumpulan prototipe yang kompak mencapai keseimbangan terbaik, dengan bank prototipe yang terlalu besar menimbulkan redundansi dan pemanfaatan yang kurang optimal, sementara terlalu sedikit prototipe gagal menangkap keragaman interaksi agen dan lingkungan. Hal ini selaras dengan teori-teori yang lebih luas tentang pembelajaran representasi, yang menekankan kegunaan primitif yang kompak dan terpisah sebagai dasar untuk generalisasi komposisional. Dalam hal ini, MoDP mendukung pandangan bahwa evolusi diri yang efektif membutuhkan bias induktif dalam bentuk pemisahan dan mekanisme arsitektur untuk spesialisasi selektif. Secara kolektif, studi ablasi ini menetapkan bahwa manfaat MoDP bukanlah kebetulan tetapi muncul dari mekanisme yang dirancang dengan cermat, yang menyeimbangkan modularitas, kekompakan, dan adaptabilitas yang ditargetkan.

Dari perspektif konseptual, MoDP memberikan wawasan berharga bagi visi yang lebih luas tentang agen yang berevolusi sendiri. Faktorisasi ruang interaksi menjadi primitif berkardinalitas rendah yang dapat dikombinasikan menawarkan strategi konkret untuk mengatasi salah satu tantangan paling sulit dalam kecerdasan buatan: generalisasi kombinatorial. Dengan mereduksi ruang masalah yang berdimensi tinggi dan saling terkait menjadi himpunan prototipe yang terstruktur dan tidak saling terkait, MoDP mengubah tantangan ekstrapolasi ke komposisi yang belum pernah dilihat sebelumnya menjadi tantangan interpolasi dalam ruang laten yang terstruktur, sehingga membuat masalah tersebut dapat diatasi. Wawasan ini selaras dengan tradisi panjang ilmu kognitif dan pembelajaran mesin yang memandang modularitas dan komposisionalitas sebagai pendorong utama generalisasi sistematis.

Sama pentingnya adalah penggabungan pemisahan representasional MoDP dengan spesialisasi tingkat arsitektur. Mekanisme ganda ini memastikan bahwa adaptasi terkonsentrasi di tempat yang memberikan manfaat terbesar untuk generalisasi sekaligus mengurangi transfer negatif antar tugas yang tidak terkait. Dengan menghubungkan struktur representasional dengan modularitas fungsional, MoDP mengilustrasikan bagaimana agen yang berevolusi sendiri dapat menghindari jebakan adaptasi monolitik, di mana pembelajaran baru sering kali mengganggu keterampilan yang telah diperoleh sebelumnya. Sebaliknya, ia

menunjukkan jalur menuju peningkatan komposisional yang berkelanjutan, di mana komponen khusus berevolusi secara terkoordinasi namun terpisah.

DAFTAR PUSTAKA

Abdurahman, A. (2026). *LITERASI BERPIKIR KRITIS TERHADAP KECERDASAN BUATAN*. NEXUS BOOKS.

Adriyendi, A. (2025). *ARTIFICIAL INTELLIGENCE DENGAN MODEL DISKRIMINATIF DAN MODEL GENERATIF (Pendekatan Saintifik Populer Pada Riset Kolaboratif)*. Penerbit Widina.

Aly, A. H. (2025). *Pesantren Digital: Masa Depan Pendidikan Islam Di Era Kecerdasan Buatan*. Publica Indonesia Utama.

Andana, E. K., & Lubis, M. (2025). *Kecerdasan Buatan Untuk Kesehatan*. Mukhlisina Revolution Center.

Anwar, S. (2026). *KURIKULUM 5.0 INTEGRASI INOVASI, TEKNOLOGI DAN KECERDASAN BELAJAR*. PT. DOA PARA WALI.

Apriadi, E. A., Julianto, R., Dwiatmoko, F., Kom, S., Kom, M., Bisri, M., & Kom, M. (2025). *KECERDASAN BUATAN Teori, Implementasi, Dan Aplikasi Di Era Digital*. Eko Aziz Apriadi.

Baskoro, S. E., Mardiah, M., Rustiyana, H. R., Nurdiyah, D., Juansa, A., Hendratni, T. W., ... & Damanik, F. H. S. (2025). *Kecerdasan Buatan Dalam Kehidupan Nyata: Konsep, Algoritma Dan Penerapan*. Star Digital Publishing.

Basyar, A. B. B., & Mukhbir, M. (2025). Evolusi Paradigma Kebijakan Pendidikan Islam Di Indonesia (1945–2025): Dari Kontrol Negara Menuju Dualisme Filosofis. *AS-SUNNIYYAH*, *5*(01), 72-83.

Benrahman, H., Kom, S., MSI, M., Puspitarini, C. E. W., MT, M., Paliling, A., ... & Prasetyo, M. A. KECERDASAN ARTIFISIAL.

Bernadus, B. B. (2025). Filsafat Teknologi Martin Heidegger Dan Artificial Intelligence: Tinjauan Kritis Terhadap Eksistensi Manusia Di Dunia. *Lumen Veritatis: Jurnal Filsafat Dan Teologi*, *16*(1), 59-70.

Dahono, Y. (2026). *Santai Aja, Ini Cuma AI: Panduan Bahasa Manusia Untuk Orang Awam (Belajar AI Pemula)*. Self Publishing.

Damastuti, F. A., Mahardika, F., Hadmar, A. A. P. M., Sanger, J. B., & Putra, Y. (2026). *KECERDASAN BUATAN DAN KOMPUTASI MODERN: KONSEP, TEKNIK, DAN APLIKASI*. Digital Publishing Library.

Dana, G. W. P., Hariyanti, N. K. D., Kumara, I. M. S., Hendrawan, A., Jati, G. P. R. S., Ardiantika, R., & Athalina, G. (2026). *Generative Artificial Intelligence: Teori, Model, Dan Implementasi*. Star Digital Publishing.

Depan, C. M., & Sebagai Mitra, A. I. (2025). Pembelajaran Mandiri. *PEMANFAATAN AI UNTUK MENGARAHKAN (AI-DIRECTED), MENDUKUNG (AI-SUPPORTED), DAN MEMBERDAYAKAN (AI-EMPOWERED) PEMBELAJARAN*, 65.

Fitroh, Z. (2026). *Cody AI: Analisis Kode Skala Besar Dan Navigasi Basis Data Kompleks*. Afdan Rojabi Publisher.

Hameida, M. K. N., Rahmadhani, D. C., & Effendi, D. N. N. (2026). *Hidup Di Era AI: Cara Berkembang Dan Bertahan Di Masa Depan Digital*. Moulin Khasya Nur Hameida.

Hananuraga, R., Afriani, G., Laksmi, N. D., Arsyad, M., & Darimis, D. (2026). Peran Artificial Intelligences Dalam Meningkatkaan Keterampilan Komunikasi Mahasiswa: Sebuah Tinjauan Literatur. *RIGGS: Journal Of Artificial Intelligence And Digital Business*, *4*(4), 2817-2824.

Helsa, Y. (2025). *Artificial Intelligence Untuk Pendidikan Strategi Pembelajaran, Efisiensi Guru, Dan Implementasi Mengajar AI Untuk Siswa Di Setiap Level*. Deepublish.

Herabadi, A. G. (2025). Media Sosial Dan AI (Artificial Intelligence): Kawan Atau Lawan? Pemberdayaan Manusia Di Era Teknologi Digital. *Perkembangan Informasi Dan Teknologi Digital Bagi Kehidupan Manusia Kini & Akan Datang*, 159.

Huda, S., Azhar, M. A. Z., Putri, N. R., Damayanti, A., Zahrani, F., Sari, D. A., ... & Qalbiyah, F. (2025). Kreasi Dakwah Di Era AI: Optimalisasi Penyiaran Islam Dengan Kecerdasan Buatan.

Huriani, Y., & Rahman, M. T. (2026). Dari Aural Ke Artificial Intelligence: Lintasan Komunikasi, Agama, Dan Isu-Isu Kontemporer.

Iskandar, I., SE, D. M. T., Min, M., Th, M., Min, D., & Bagoes, A. (2026). *TRANSFORMASI PENDIDIKAN DI ERA KECERDASAN ARTIFISIAL*. Goresan Pena.

Kamuri, K. J. (2026). *Yang Disingkirkan Oleh AI (Artificial Intelligence)*. CV. Vocezmi Learnov.

Kariono, M. S. (2026). *MEMANUSIAKAN TEKNOLOGI: PERJALANAN KESADARAN DIRI DI ERA KECERDASAN BUATAN*. Dunia Penerbitan Buku.

Kesumaningtyas, F., Purnama, P. A. W., Indahsari, A. N., & Pratama, R. (2025). *Revolusi Kecerdasan Buatan: Mengubah Masa Depan Teknologi*. Serasi Media Teknologi.

Kom, A. Z. M., & Fordiana Ekawati, S. E. (2025). *Penerapan Artificial Intelligence (AI) Pada Peralatan Militer*. CV Eureka Media Aksara.

Kurniawan, D., & Rojabi, M. A. (2026). *Halo AI: Revolusi Asisten Virtual Berbasis Kecerdasan Buatan*. Afdan Rojabi Publisher.

Kushariyadi, K., Apriyanto, H., Herdiana, Y., Asy'ari, F. H., Judijanto, L., Pasrun, Y. P., & Mardikawati, B. (2024). *Artificial Intelligence: Dinamika Perkembangan AI Beserta Penerapannya*. PT. Sonpedia Publishing Indonesia.

Luckianto, M., Kom, S., Kom, K. S., Kom, M., Utomo, Y., & Kom, S. (2026). *The Rise Of Agentic Intelligence: Transformasi Menuju Kecerdasan Otonom Di Era Kecerdasan Buatan Modern*. Binus Media Publishing (PT. Widia Inovasi Nusantara).

Masmur, M. (2025). Transformasi Manajemen Dosen Pendidikan Tinggi Melalui Integrasi Artificial Intelligence Dalam Pembelajaran. *Jurnal Ilmiah Edukatif*, *11*(2), 415-434.

Mumu, A. S., Th, S., & Tamaweol, R. D. (2025). *Teologi Kecerdasan Buatan (AI): Tinjauan Dogmatis, Praktis, Dan Reflektif*. PT. Revormasi Jangkar Philosophia.

Muvid, M. B., Fendy Hariatama, H., Baskara, F. R., Firmansyah, M. P. I., Rakhman, A. S., Sutomo, E., ... & Mashuri, K. (2025). Mendesain Pembelajaran Humanistik Untuk Membentuk Self-Directed Learning Bagi Mahasiswa. *Repository Global Aksara Pers*, *5*(1).

Nampira, A. A., Judijanto, L., Wati, D. C., Hermawan, E., Cahyono, T. A., Prayudani, S., ... & Sitanggang, A. T. (2025). *Artificial Intelligence: A Guide For Thinking Humans*. PT. Green Pustaka Indonesia.

Nugraha, D. K. S., & Kesa, D. D. (2025). Kreativitas Berbasis Data: Pemanfaatan Kecerdasan Buatan Dalam Inovasi Pemasaran Untuk Ketahanan Ekonomi. *KALBISOCIO Jurnal Bisnis Dan Komunikasi*, *12*(3), 108-119.

Nurdiana, I. (2026). *Jan AI: Mendefinisikan Ulang Privasi Dalam Ekosistem Kecerdasan Buatan*. Afdan Rojabi Publisher.

Nurmiati, E., & Azzahra, A. S. N. (2026). CIA Triad In The Age Of AI: Tinjauan Etika Profesi Atas Konflik Anthropic–Pentagon. *Jurnal Surya Informatika*, *16*(1), 145-151.

Parrangan, J., Mappanyompa, M., Trisnawan, A. B., Dongka, R. H., Kadir, Y., Wiranto, I., ... & Putra, I. K. A. (2026). *Pengantar Artificial Intelligence (AI)*. CV. Edu Akademi.

Permana, Y. P., Kom, S., & Kom, M. (2025). *Kecerdasan Buatan: Mengubah Dunia Dengan Teknologi*. Mega Press Nusantara.

Puspita, A. F., SE, M., Rusydi, M. K., Ak, M., Hendi Subandi, S. E., & Nurkholis, P. D. (2025). *MASA DEPAN Taxing Artificial Intelligence Dan Implikasi Etis*. PT. Rajagrafindo Persada-Rajawali Pers.

Putra, R. P., Rasyid, R. W., Usman, S., & Nurfadillah, N. (2025). *Guru Dan Mesin: Antara Hikmah Dan Kecerdasan Buatan*. CV Eureka Media Aksara.

Ramdhon, A. N., Kom, S., & Kom, M. (2026). *SINYAL OTAK DAN BARIS KODE: Revolusi Informatika Di Era Kecerdasan Buatan*. PT KIMHSAFI ALUNG CIPTA.

Rianto, S., Yudianto, D., Yuana, F., & Dardiri, A. A. (2024). *Penelitian Berdaya AI: AI Senjata Rahasia Baru Ilmuwan Merubah Dunia*. Media Nusa Creative (MNC Publishing).

Rijal, S., AP, M., Jusman, S., AP, M., Suleman, S., Sos, S., ... & Amiruddin, S. E. (2026). *KEBIJAKAN PUBLIK 5.0:(TRANSFORMASI DIGITAL PEMERINTAHAN DI ERA SOCIETY 5.0)*. Dunia Penerbitan Buku.

Rohmah, L., & Hilalludin, H. (2025). Ai Tanpa Tuan: Revolusi Otonomi Algoritma Dalam Mengambil Keputusan. *An-Nuriyah: Journal Of Islamic Technology And Informatics Education*, *1*(04), 23-35.

RR Ella Evrita H SE, M. M. (2025). *Digital Darwinism: Hukum, Kreativitas, Dan Evolusi Media Di Era AI*. PT Indonesia Delapan Kreasi Nusa.

Rudi Nopiansyah, S. H., & MH, C. (2025). *HUKUM DAN KECERDASAN BUATAN: Menyongsong Era Baru Dunia Hukum*. PENERBIT KBM INDONESIA.

Salsabila, O. N. (2026). *Masa Depan Personal AI: Integrasi Kecerdasan Dalam Aktivitas Keseharian*. Afdan Rojabi Publisher.

Santoso, H. I. (2026). *Intelligent Logistics: Pemanfaatan Data Analytics Dan Agen AI Dalam Pengambilan Keputusan*. Hertadi Iman.

Siswanto, D. H. (2026). *Heutagogi Dan Inovasi Pendidikan: Evolusi Paradigma Pembelajaran Kontemporer*. Lenggo Geni Pustaka.

Solihin, R., Sos, S., & Han, M. (2026). *QUANTUMSHADOW: INTELIJEN DAN PERANG DI ERA MULTIPOLAR*. Penerbit Widina.

Sutopo, A. H. (2026). *Moralitas Digital: Rekonstruksi Pendidikan Moral Di Era AI Dan Masyarakat Algoritmik*. Topazart.

Thomas, S. T., & Santoso, E. H. (2025). *MESIN YANG BERPIKIR, DATA YANG BERCERITA: Revolusi Artificial Intelligence Dalam Dunia*. Global Kreatif Media.

Tonggiroh, M., Sumantyo, F. D. S., Prilosadosa, B. H., Ds, M., Vandika, A. Y., Sumitra, T., ... & Kom, M. Artificial Intelligence: Konsep Dan Penerapan Pada Kehidupan Modern.

Tontowi, A. E. (2025). *Transformasi Digital, Harmonisasi Inovasi, Dan Kemanusiaan*. UGM PRESS.

Tradesa, A. P., & Alfarizi, L. M. (2025). Status Hukum Karya Intelektual Yang Dihasilkan Oleh Kecerdasan Buatan Dalam Rezim Hukum Paten Indonesia.

Triana, Y. S., Kom, M., Hadjon, R. P., Kom, S., Kom, M., Salamah, U., & Fajri, M. (2026). *ETIKA DI ERA KECERDASAN BUATAN*. Penamuda Media.

Utari, P. (2026). *DINAMIKA PERKEMBANGAN TEORI KOMUNIKASI Evolusi, Konvergensi, Teknologi Dan Dekolonisasi*. PENERBIT KBM INDONESIA.

Wardani, F. P. (2025). *SINGULARITY TECHNOLOGY "Ketika Teknologi Melampaui Batas Dan Tidak Bisa Dikendalikan Lagi"*. Feniks Muda Sejahtera.

Wardhana, D. A. P., ST, S., Hardiansyah, B., Kom, S., Hartono, I. E. D., Mandita, F., ... & ST, S. (2025). *Dasar-Dasar Artificial Intelligence: Pendekatan Praktis Dengan Python Dan C++*. PT Bukuloka Literasi Bangsa.

Wibisono, G., Raharya, N., & Prasetio, S. (2026). *Artificial Intelegence Dalam Transformasi Bisnis Modern*. Deepublish.

Wibowo, I. A., & Kom, M. (2025). *Artificial Intelligence (AI) Dalam Jaringan Otonom*. Yayasan Penerbit.

Wijanarko, S., Ekaningsih, L., & Farida, A. (2026). Konstruksi Pertanggungjawaban Pidana Korporasi Pengembang Kecerdasan Buatan (AI) Terhadap Tindak Pidana Pemalsuan Dokumen Elektronik. *Lex Journal: Kajian Hukum Dan Keadilan*, *10*(2), 329-341.

Yazid, A. (2025). *SDM 5.0: Strategi Manusia Dalam Era AI Dan Otomatisasi*. CV Eureka Media Aksara.

Yusnaini, Y., Muhaimin, M., Firmansyah, F., Setiwan, Y. L., Bakhtiar, R., Aisyah, A., ... & Suhaeri, S. (2024). *Artificial Intelligence Dalam Perkembangan Teknologi Komunikasi*. CV. Gita Lentera.

EVOLUSI AI OTONOM YANG BERWUJUD

Dr. Joseph Teguh Santoso, S.Kom, M.Kom.

BIODATA PENULIS

Dr. Joseph Teguh Santoso, M.Kom memiliki Jabatan Akademik Lektor Kepala dan praktisi industri yang berpengalaman. Saat ini menjabat sebagai Rektor Universitas Sains dan Teknologi Komputer (Universitas STEKOM), salah satu universitas terkemuka di Jawa Tengah, Indonesia. Dengan pengalaman lebih dari 13 tahun di dunia bisnis dan praktisi industri di China, beliau membawa perspektif global dan inovasi yang signifikan ke dalam dunia akademis.Sebagai seorang entrepreneur, penulis adalah pencipta TopLoker.com, sebuah platform inovatif yang merevolusi cara mencari dan menawarkan pekerjaan. TopLoker.com adalah portal lowongan bursa kerja terbesar di Indonesia, khusus untuk pendidikan SMA/SMK sederajat. TopLoker.com telah mendapatkan penghargaan sebagai juara 1 Startup4industry 2022 oleh Kementerian Perindustrian Republik Indonesia. Kontribusi Dr. Joseph dalam menyediakan akses pekerjaan yang luas bagi lulusan SMA/SMK telah membantu banyak individu menemukan peluang kerja yang sesuai dengan keahlian mereka. Selain itu, Dr. Joseph Teguh Santoso, M.Kom adalah pendiri dari dua organisasi yaitu (1) organisasi guru/pendidik PTIC (Perkumpulan Teacherpreneur Indonesia Cerdas) yang bertujuan untuk meningkatkan kualitas pendidikan dan kesejahteraan guru/pendidik dengan wawasan entrepreneurship, serta (2) organisasi industri PERKIVI (Perkumpulan Komunitas Industri dan Vokasi Indonesia) yang berfokus pada pengembangan link and match antara industri dan dunia pendidikan. Sebagai Rektor, Dr. Joseph Teguh Santoso, M.Kom memiliki kepemimpinan yang berorientasi pada hasil, dan berkomitmen untuk mendorong kemajuan Universitas Sains dan Teknologi Komputer (Universitas STEKOM). Saat ini Universitas STEKOM telah mengalami transformasi positif dalam peningkatan kualitas pendidikan, perluasan fasilitas, serta penguatan kemitraan Perguruan Tinggi Nasional dan Internasional. Beliau memprioritaskan pengembangan sumber daya manusia dan penelitian, serta memastikan bahwa universitas berada di garis depan dalam inovasi dan teknologi untuk mencapai tujuan akhir, yaitu lulusan yang mampu bekerja dan sukses setelah lulus. Dr. Joseph Teguh Santoso, M.Kom sering diundang sebagai pembicara di berbagai konferensi nasional maupun internasional dan telah menerima berbagai penghargaan atas dedikasinya dalam bidang pendidikan, industri, dan kewirausahaan.

YAYASAN PRIMA AGUS TEKNIK

PENERBIT :
YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7695-42-0 (PDF)
9 786347 695420