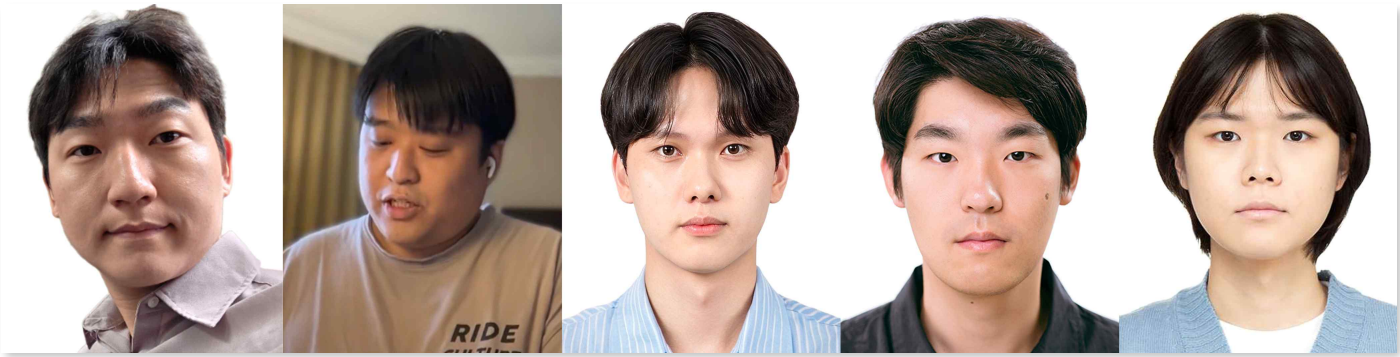
GIST developed an AI that understands human-like reasoning: Accuracy improved by 5.85%p compared to the existing model (83.59% → 89.44%)

- Professor Sundong Kim's team from the Department of AI Convergence proposes an "intention learning algorithm" that estimates and sorts human intentions during problem solving, and develops a "data augmentation technique" that generates diverse solution paths.

- By combining alignment and augmentation, the team demonstrates the potential for "human-like reasoning" beyond simple correct answers... Presented orally at ACM SIGKDD 2025, the premier data science conference, and published in the international journal 《TMLR》



▲ Professor Sundong Kim of the Department of AI Convergence at GIST (corresponding author of the KDD and TMLR papers), postdoctoral researcher Sejin Kim (first author of the KDD paper and co-corresponding author of the TMLR paper), master's student Sanha Hwang (first author of the TMLR paper), master's student Seungpil Lee (second author of the TMLR paper), and undergraduate student Ho-seong Lee of the Department of Electrical Engineering and Computer Science (second author of the KDD paper)

The Gwangju Institute of Science and Technology (GIST, President Kichul Lim) announced that a research team led by Professor Sundong Kim of the Department of AI Convergence has proposed a data augmentation technique that combines a learning algorithm that estimates and organizes the "intentions" embedded in human problem-solving processes with an AI generative model* to generate diverse solution processes, much like humans.

By combining these two approaches, the research team demonstrated that AI can achieve "human-like reasoning abilities*" beyond simply deriving correct answers.

* generative model: An AI model that generates new data (e.g., images, text, and problem-solving processes) rather than simply classifying or predicting existing data. Representative examples include GANs, VAEs, large-scale language models (LLMs), and GFlowNet.

* solution augmentation: A method of generating diverse solution processes from existing data and using them as training data. This study utilized GFlowNet for implementation.

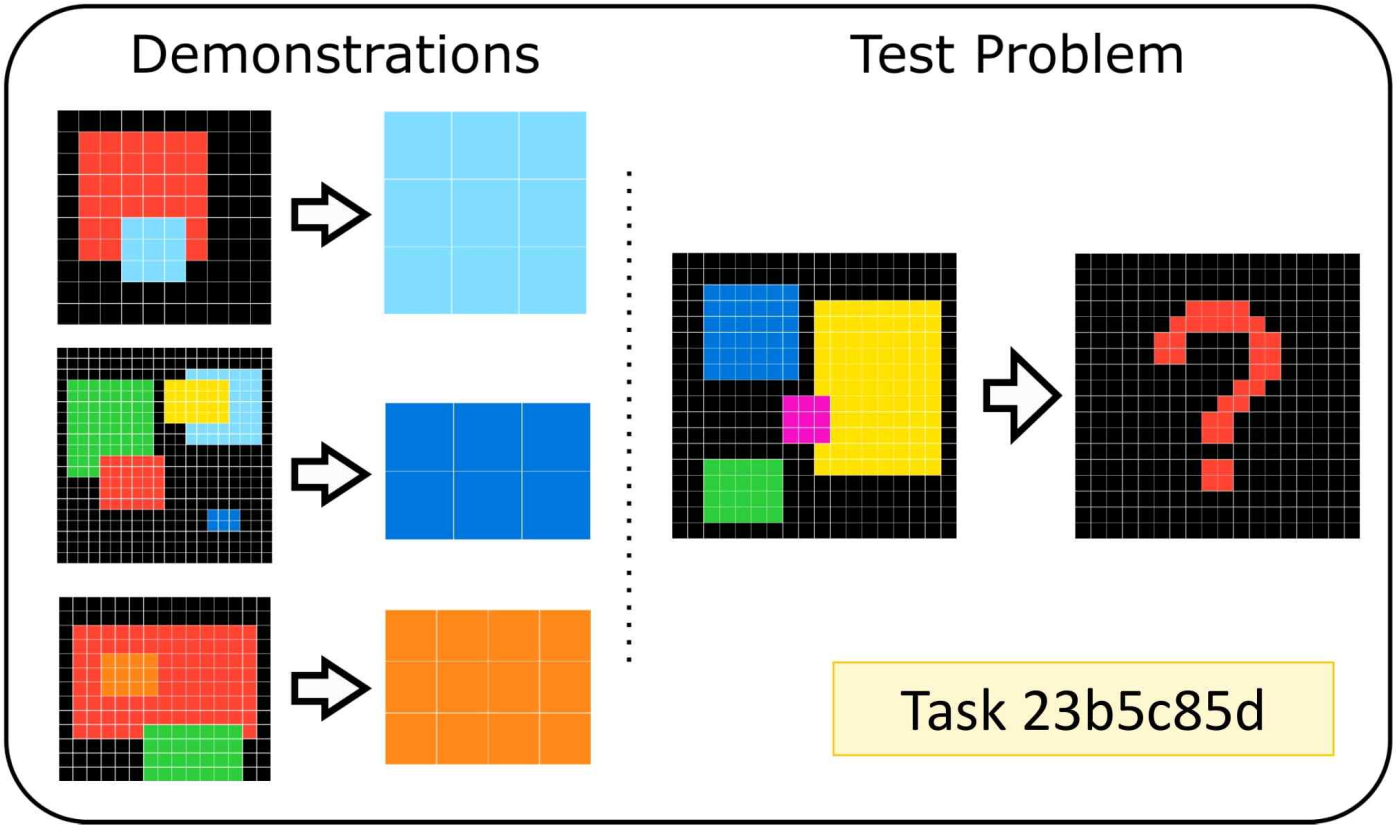
* human-like reasoning: The ability to arrive at answers through a human-like thought process, rather than through simple calculations.

Today, artificial intelligence demonstrates its strength in quickly deriving answers to given problems, but it still lacks the reasoning ability to solve problems through a step-by-step thought process like humans.

Humans experience trial and error when solving problems, attempting the same goal in various ways. The accumulated solution data during this process is not a simple list of actions, but rather contains intentions*. Learning these intentions is the core of AI possessing "human-like reasoning."

* intention: The ultimate goal or strategy a person seeks to achieve during a problem-solving process. The same intention can manifest itself in multiple behavioral patterns.

From this perspective, the research team developed an algorithm to estimate and sort intentions during the problem-solving process* and proposed a data augmentation technique that generates diverse solution trajectories using GFlowNet*, a generative model.



▲ An example problem for collecting trajectory data containing the solution process. This visualization illustrates one of the ARC-AGI benchmark problems, which requires identifying the problem's rules from the pairs of images on the left and completing the pair of images on the right. The rule for this problem is to select the rectangle with the smallest area from the left image and place it on the right. By reflecting this, we implemented learning that mimics the human thought process.

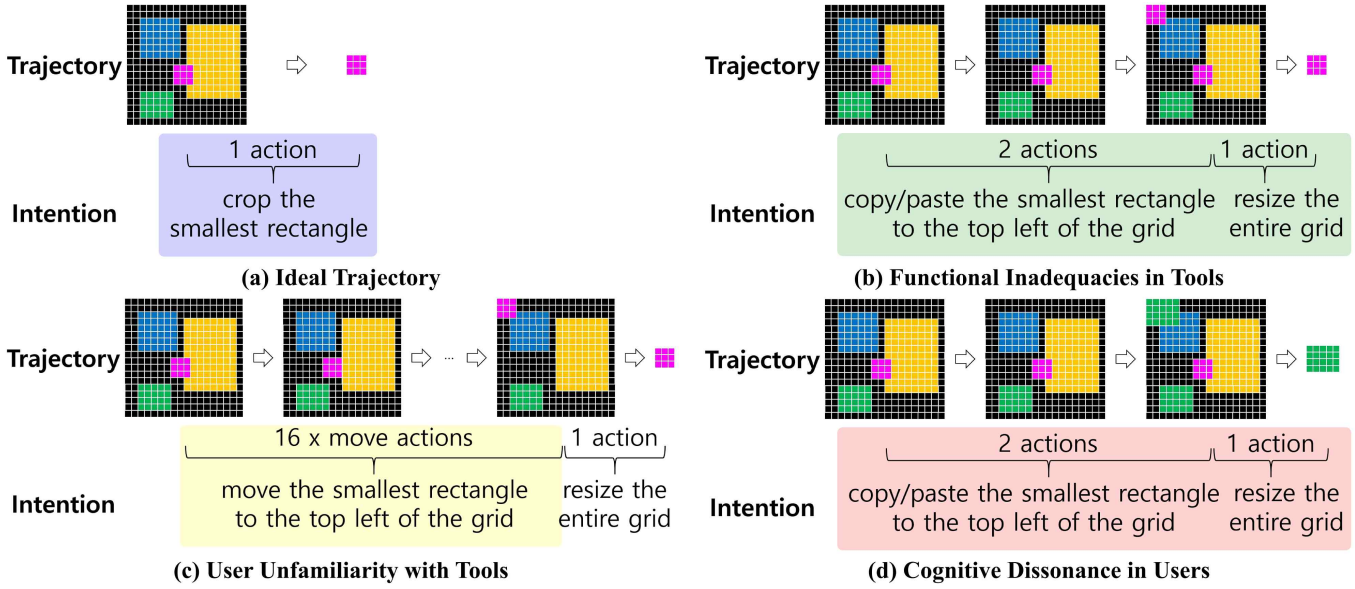
The former reflects the human thought process, while the latter expands the solution process to enhance generalization performance. The two approaches complement each other. Through this, the goal was to develop an AI model that goes beyond simply deriving the correct answer and possesses human-like thinking and generalization capabilities.

* alignment: The process of aligning the AI's estimated intent with the actual problem-solving process.

* GFlowNet (Generative Flow Network): A generative modeling technique that generates a variety of possible solutions through probabilistic exploration.

The research team analyzed human problem-solving data and categorized cases where optimal solutions were not reached into three categories (① insufficient functionality, ② inefficient attempts, and ③ incorrect strategies), incorporating these into learning.

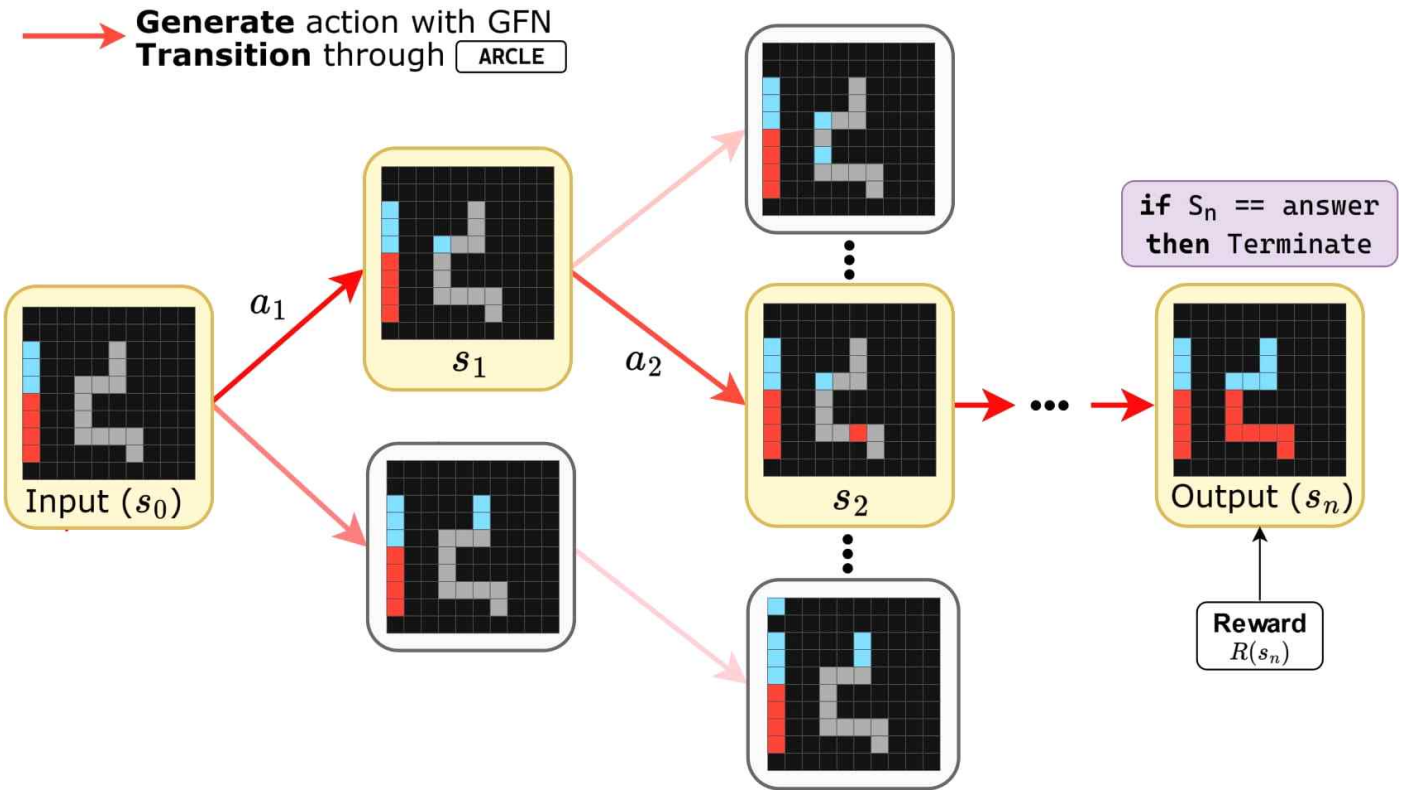
Next, they proposed an algorithm that divides the problem-solving trajectory into multiple stages and estimates and aligns the intent at each stage. By incorporating this into AI learning, they implemented learning that mimics the human thought process.



▲ Examples of various problem-solving trajectories and intentions. This figure illustrates the various trajectories that a person can take when solving the same problem, along with the underlying intentions. (a) Ideal Trajectory: The most efficient solution that achieves the goal with the fewest actions. (b) Functional Inadequacies in Tools: Unnecessary steps added due to limitations in available functions. (c) User Unfamiliarity with Tools: Unnecessary steps added due to lack of familiarity with the tools. (d) Cognitive Dissonance in Users: Unnecessary steps taken due to the user's choice of a different strategy, resulting in a discrepancy between the user's intention and the actual solution process.

Furthermore, by using a data augmentation technique based on ZipflowNet, we significantly increased the diversity of training data and generalization performance by generating diverse solution processes, much like humans.

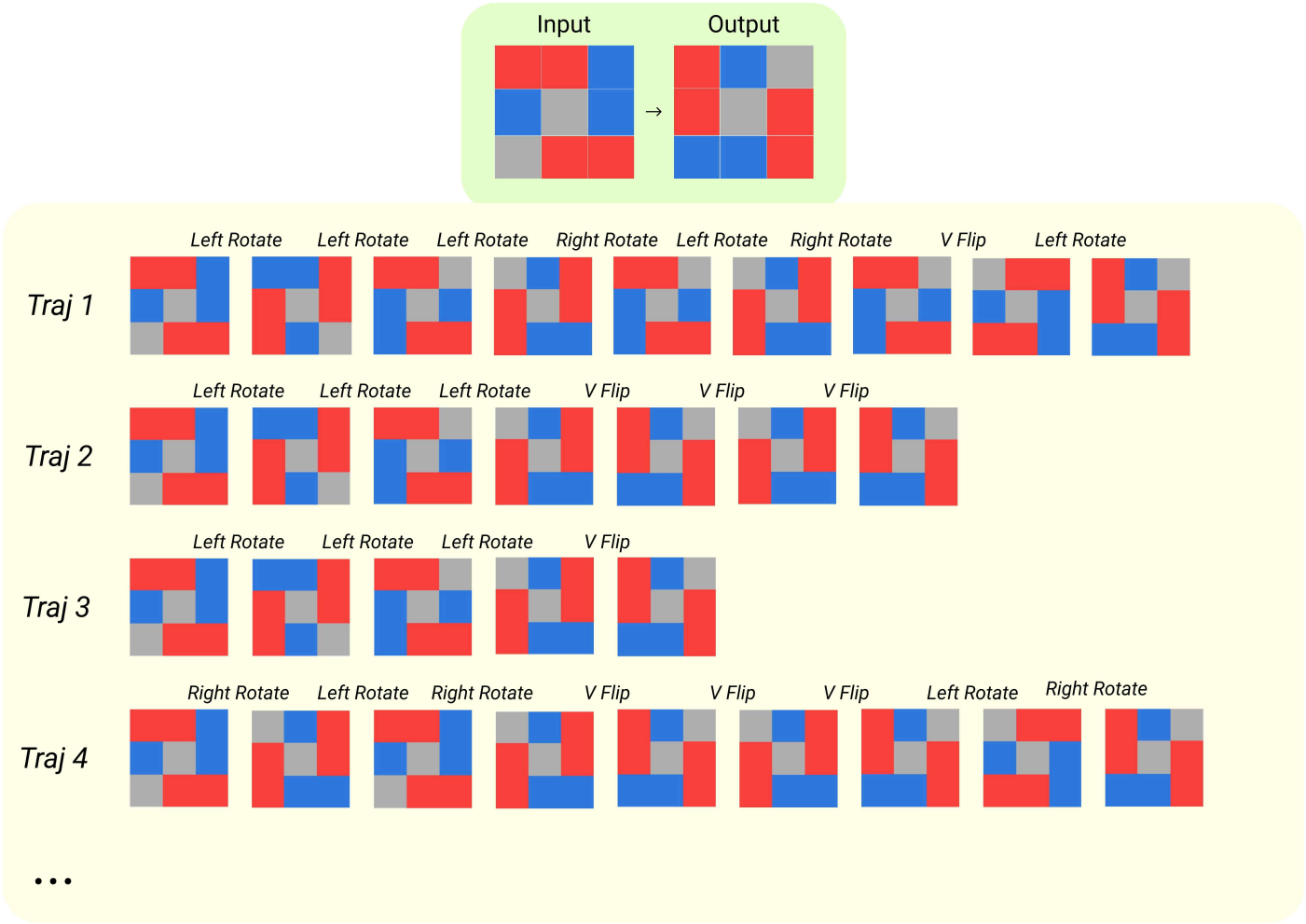
ZzipflowNet generates multiple solution paths through probabilistic exploration, passing through multiple intermediate states. Because multiple trajectories are generated for the same input using various combinations of operations, such as rotation and symmetry, and utilized for learning, it can learn diverse strategy distributions without being biased toward a specific solution method.



▲ The problem-solving trajectory generation process based on ZipflowNet. ZipflowNet starts from an input state and explores various intermediate states through a series of actions, ultimately reaching the correct output. This probabilistic exploration process allows it to generate multiple solution paths, much like a human, significantly expanding the diversity of training data, contributing to improved generalization performance.

The research team collected solution trajectory data from real people and, when trajectories were insufficient, generated and supplemented them with GFlowNet.

Using this diverse solution data for training, the accuracy improved by approximately 5.85 percentage points (83.59% → 89.44%) compared to the existing model. This demonstrates that AI can reason and generalize like a human.



▲ Various solution trajectories generated for the same diagonal flip problem. A single input-output pair can undergo multiple transformation operations (rotation, symmetry, etc.) to arrive at the same correct answer. By generating these diverse solution processes with GFlowNet and using them as training data, the research team empowered the AI to consider multiple strategies, just like a human.

Professor Sundong Kim stated, "Humans often find solutions cleverly using familiar methods rather than following standard problem-solving processes. However, in AI model training, human-collected data is often used without much thought." He continued, "Systematically applying the pre- and post-processing processes presented in this paper will allow us to overcome inherent limitations of data and develop AI that exhibits more desirable behaviors."

He added, "Furthermore, the data generation technique presented in this study will significantly contribute to expanding training data in fields where human-solution data is insufficient."

Dr. Sejin Kim stated, "This research is the first step toward creating AI that goes beyond simply guessing the correct answer; it can understand human problem-solving intentions and autonomously learn various solution processes. This research could be expanded into creative problem-solving AI, collaborative AI for education, and AI that adapts to new environments."

The research project, “Development of an Algorithm for Estimating Intentions Embedded in Human Problem-Solving Processes,” supervised by Professor Sundong Kim of the Department of AI Convergence at GIST and conducted by postdoctoral researcher Sejin Kim and undergraduate student Ho-seong Lee, was supported by the Information and Communications Technology Development Project, the International Joint Research Project on Digital Innovation Technology, and the Graduate School of Artificial Intelligence Support Project of the Institute of Information & Communications Technology Planning and Evaluation (IITP) under the Ministry of Science and ICT, the Mid-Career Researcher Support Project and the Postdoctoral Domestic Training Support Project of the National Research Foundation of Korea (NRF), and the GIST Postdoctoral Researcher Value-up Program. The "GFlowNet-based Problem Solving Trajectory Augmentation Study," conducted by Sanha Hwang (graduated), Seungpil Lee (graduated), and Sejin Kim (postdoctoral researcher) in the Department of AI Convergence, was supported by the Information and Communications Technology Development Project and the International Joint Research Program on Digital Innovation Technology of the Institute of Information and Communications Planning (IITP) under the Ministry of Science and ICT, the National Research Foundation of Korea (NRF)'s Mid-Career Researcher Support Program and Postdoctoral Domestic Training Support Program, and the GIST Future-Leading Specialized Research Program.

The research results were presented orally on August 6th at ACM SIGKDD* 2025, a world-renowned academic conference in data science, and published in the September 2025 issue of the international journal 《Transactions on Machine Learning Research (TMLR*)》.

* ACM SIGKDD (Conference on Knowledge Discovery & Data Mining): Hosted by the ACM's Special Interest Group on Data Mining and Knowledge Discovery (SIGKDD), this world-class data science conference presents the latest research findings and industrial applications in artificial intelligence, machine learning, and big data analytics.

* Transactions on Machine Learning Research (TMLR): Launched to rapidly share research results in the field of machine learning, the journal is supported by the Society for Neural Information Processing Systems (NeurIPS). It promotes transparency and active exchange among researchers through a faster peer review process than traditional journals, and covers a wide range of topics, from machine learning theory to applications.

Meanwhile, GIST stated that this research achievement considered both academic significance and industrial applicability, and that discussions regarding technology transfer can be made through the Technology Commercialization Center (hgmoon@gist.ac.kr).

GIST
Since 1993