

News Release

FOR IMMEDIATE RELEASE

Development of business-oriented AI technology based on competitive self-play learning

Confirmed that loss from over or under-stocking, a supply chain management challenge, can be reduced to 1/4 that of human judgment

Tokyo, December 25, 2017 --- Hitachi, Ltd. (TSE: 6501, Hitachi) today announced the development of business-oriented AI (artificial intelligence) technology that learns by representing business as groups of interconnected AI and having the AI groups compete among themselves on a computer without relying on actual human-prepared data. When this AI technology was applied to the "beer distribution game" that simulates a supply chain consisting of multiple companies, the results confirmed that it was possible to reduce loss from over or under stocking to about 1/4 that of decisions based on human experience. The effectiveness of AI learning from self-play has already been demonstrated for competitive games such as Go. This new development indicates that learning based on competitive self-play is also effective for business problems with many uncertain factors.

In general, AI that uses machine learning such as deep learning, predicts and makes decision based on what it has learnt from a large set of actual data. Thus there is the issue that when a large set of actual data is not available, it becomes difficult to make accurate predictions and decisions. While it has been shown that AI can effectively learn from a large volume of self-generated data by repeatedly playing against itself in competitive games such as Go, it was unclear whether this could be applied to business problems with many uncertain factors.⁽¹⁾

This time, Hitachi has developed AI technology that utilizes competitive self-play for learning and is applicable to business problems. In the AI technology developed, companies involved in a business are represented by AI agents using deep learning, and a business is represented by an AI group in which multiple AI agents are interconnected. Each AI agent learns effective actions to improve a given outcome⁽²⁾ such as reducing loss, while repetitively exchanging information and goods with other agents. Learning is executed in parallel with several AI groups generated on a computer. By repetitively performing "competitive self-play" several thousand times, the AI groups compete to achieve the best outcome among themselves (Fig. 1).

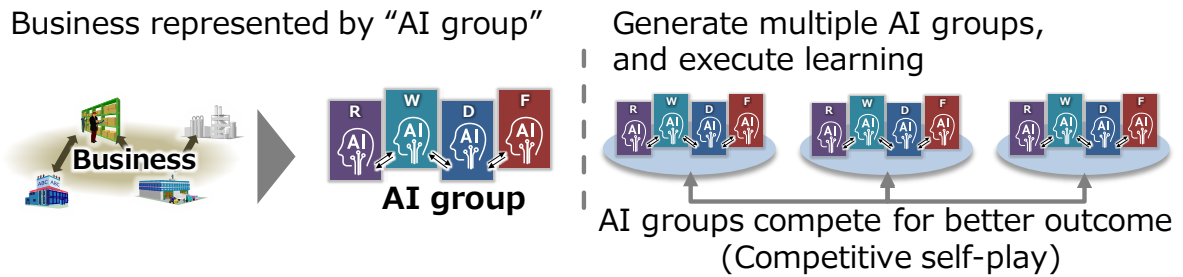


Figure 1. Learning and competitive self-play with multiple AI groups

This AI technology has the following features to realize a better outcome.

1. Improve the overall outcome of the AI group by controlling AI agent learning using a learning manager

The AI technology developed is equipped with a learning management function that manages the learning of each interconnected AI agent within the same group, and prevents the learning of one AI agent from adversely affecting another agent. This function is responsible for controlling the timing of learning of each AI agent. In the beginning, the learning manager only allows one AI agent to learn, and then gradually increases the number of learning AI agents (Fig. 2). By doing this, conflict arising from competition when AI agents learn at the same time can be avoided, and collaboration between AI agents can be learnt, leading to an improved outcome for the AI group as a whole.

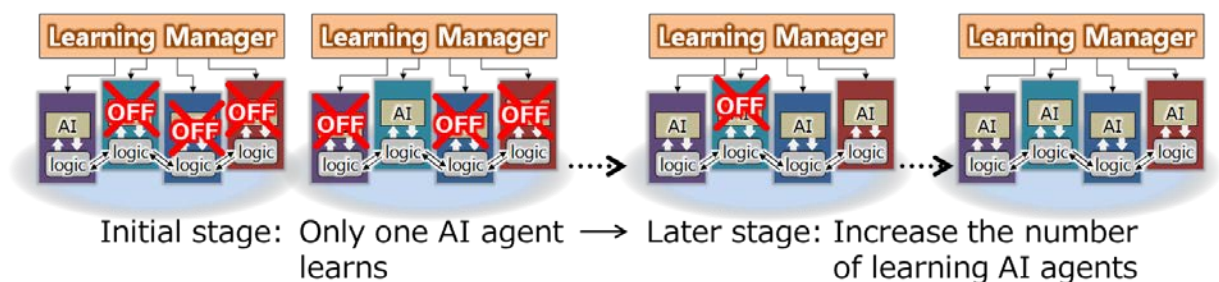


Figure 2. Learning management function that learns cooperation between AI agents

2. Function to evolve AI agents by generating better models by crossing learnt models

When AI agents within an AI group learn repeatedly, a phenomenon of where the improvements in outcome of the group may become stagnated appears as each AI agent's learning result (model) becomes biased towards the individual optimum state.

To address this issue, a function was developed to mix (cross-over) the parameters of an AI agent's model with that of other agents in the multiple AI groups on the computer, create new AI agents with new crossed models, and construct new AI groups (Fig. 3). The function then compares the outcome of the multiple AI groups generated, including that of the newly constructed AI group, keeps high performance groups while deleting poor performance groups and repeatedly executes this process (competitive self-play) to pursue better outcomes. (Fig. 4)

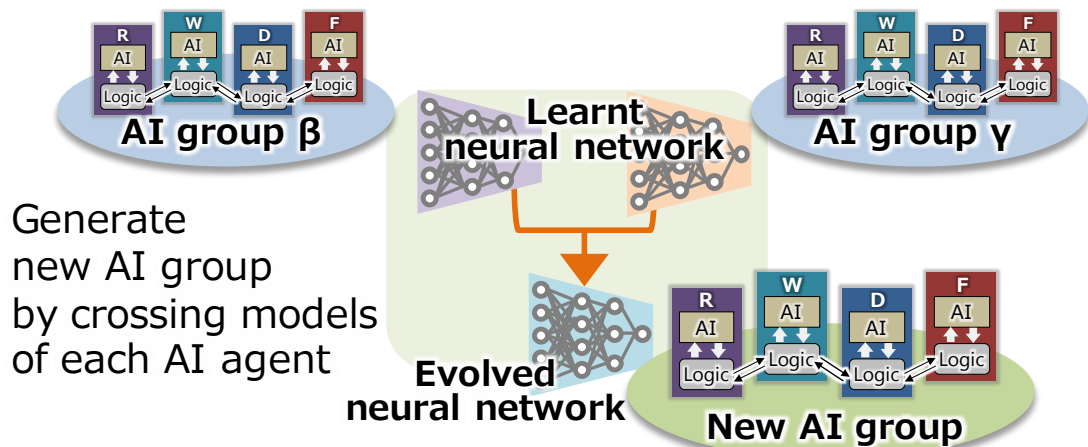


Figure 3. Technology to evolve AI agents

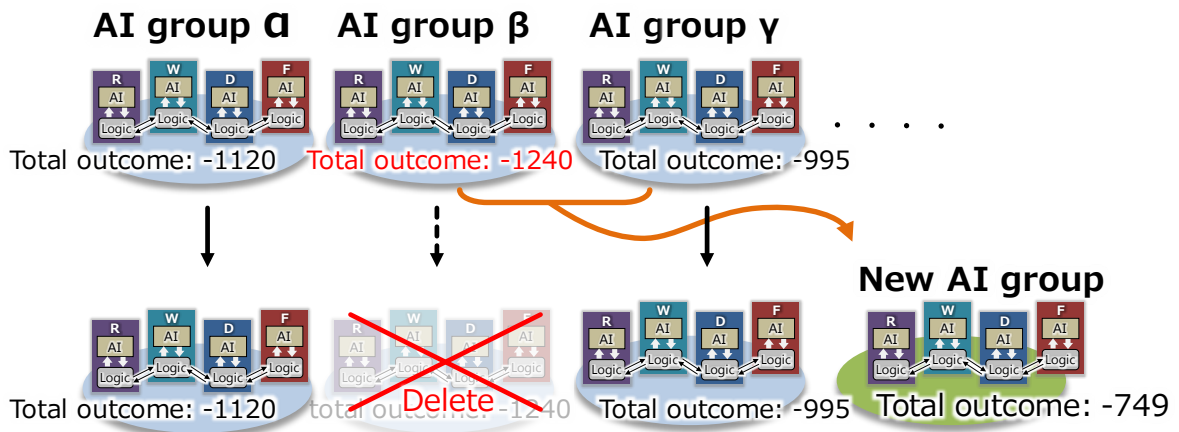


Figure 4. Competitive self-play by AI groups

The effectiveness of the AI technology developed was verified by application to the "beer distribution game" which simulates the business of multiple companies on a supply chain. In this game, four independent agents for retailer (R), wholesaler (W), distributor (D) and factory (F), decide on order volume while competing to minimize

loss in the entire supply chain from excess inventory or understocking. In addition to constantly being subjected to unpredictable demand fluctuations, a constraint of not being able to share inventory information with each other is placed on each agent, which is different from other games such as Go where all players are jointly aware of the situation. In a study where human experts played the beer game, and placed orders based on their experience, it was reported that an average of 2,028 dollar loss was incurred over 35 weeks.⁽³⁾ On the other hand, using this AI technology, it was confirmed that the loss could be reduced to 489 dollars.⁽⁴⁾ This result demonstrates that AI learning based on competitive self-play is effective in business problems as well.

The source code of the AI technology developed will be shared within the Hitachi Group so that it may be utilized in Hitachi's services and products worldwide, and thereby facilitate its use in a wide range of Social Innovation Business, such as in the power and energy, industry and distribution, water, urban, finance, government and public, and healthcare sectors.

A part of this research will be presented at the "Hawaii International Conference on System Sciences" to be held from 3-6 January 2018 in Hawaii, USA.

- (1) Research is being conducted in reinforcement learning where the AI continues to learn from not only actual data previously collected but also from the new data generated by its own actions. The application of reinforcement learning to explore actions in real world cases such as business have been restricted due to concerns of possible impact and the question of responsibility.
- (2) Outcome corresponds to a numerical value that is to be improved (maximize or minimize) in business, and is set by a human operator according to the problem.
- (3) Each agent places an order once a week without exchanging information. Loss from excess inventory or understocking is set at \$0.50 and \$1.00 per item, respectively, to calculate loss. J. D. Sterman, "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment", *Management Science*, Vol. 35, No. 3, pp. 321-339, 1989.
- (4) As was the restriction in the original beer game where the players were human, the AI agents do not share information with each other.

About Hitachi, Ltd.

Hitachi, Ltd. (TSE: 6501), headquartered in Tokyo, Japan, delivers innovations that answer society's challenges. The company's consolidated revenues for fiscal 2016 (ended March 31, 2017) totaled 9,162.2 billion yen (\$81.8 billion). The Hitachi Group is a global leader in the Social Innovation Business, and it has approximately 304,000 employees worldwide. Through collaborative creation, Hitachi is providing solutions to customers in a broad range of sectors, including Power / Energy, Industry / Distribution / Water, Urban Development, and Finance / Government & Public / Healthcare. For more information on Hitachi, please visit the company's website at <http://www.hitachi.com>.

###

Information contained in this news release is current as of the date of the press announcement, but may be subject to change without prior notice.
