May 2019 Issue

Briefing Memo

The 2nd AI Boom and its Applications for Military Systems
——Decision Support Trials and Limits of Expert Systems——

ONO Keishi , Chief, Security Studies Department

The use of artificial intelligence (AI) in human society has two main courses: improving the autonomy of machines connected via robotics technology, and greatly complementing and substituting human intellectual labor. Applying this to the part of society that is the military would result in weapon automation and decision support for commanders and field personnel, and AI has been applied to military systems along these lines. This article looks back on attempts to build AI-using military decision support systems, along with their limitations, and is centered on the expert systems that were the core of the 2nd AI boom in the 1970s-80s.

1. The 2nd AI Boom and “Expert Systems”

Development of computers is a prerequisite for the realization of AI, and AI development to date has generally been divided into three stages (see Table). The 1st AI boom began in the 1950s and was primarily aimed at searching for optimal solutions under defined conditions. The 2nd AI boom began in the 1980s. 1st boom AI “processed facts,” but 2nd boom AI “inferred from knowledge.” Finally, the 3rd AI boom has been continuing since 2010, and is driven by the evolution of machine learning (deep learning) and the use of big data.

Amidst this trend, AI since the 2nd boom has started to demonstrate practical utility. In the initial development of AI, there was a tendency for it to be in charge of solving a wide range of general problems. However, general principles cannot be easily derived from specific problems, and, even if they can be derived, cannot be applied to real problems. As such, the 2nd AI boom was aimed at the development of dedicated systems for specific issues, instead of aiming for generality.

Table: Overview of Each AI Boom

		Application Scope	Degree of Logical Dependence
1st Boom (1950s-60s)	Logic Deductive reasoning	Limited (puzzles, games, etc.)	High
2nd Boom (1970s-80s)	Knowledge Deductive reasoning	Wide (“expert systems,” etc.)	Medium
3rd Boom (2010-)	Statistics (learning) Inductive reasoning	Very wide (pattern recognition, machine translation, etc.)	Low

Source: Nishigaki Toru, *Biggu Deta to Jinkouchinou* [Big Data and Artificial Intelligence] (Chuokoron-Shinsha, Inc., 2016), Table on pg. 172, partially revised.

The 2nd AI boom is represented by “expert systems.” These are an attempt to replace advice and judgement by experts, and are largely composed of a “knowledge base” and an “inference engine.” The knowledge base is a collection of knowledge possessed by experts in a specific field, and an inference engine is a programmatization of how knowledge bases are used by experts (search procedures, criteria, etc.). While repeatedly executing logical conditions in the form of “if/then” or “yes/no” for the knowledge base in a specific field of expertise, the data is narrowed down and conclusions are drawn (deductive reasoning). MYCIN, a system developed at Stanford University in the early 1970s, had a 69% probability of accurately determining an infectious disease diagnosis (approaching the 80% probability of human experts).

2. Tactical Decision Support via Expert Systems

Expert systems, which were deemed to have reached practical use in the civilian sector, were reviewed for application to military systems. In the 1970s, the number and types of sensors increased due to the increased functionality of equipment, there was a dramatic increase in the amount of information to be processed, in addition to the information becoming more complex, and it became necessary to respond quickly during battles. Efforts were made to develop expert systems for military systems in order to overcome the physical limitations of time, as well as quality and volume of information.

For example, the CLASSIFY expert system was developed to assist in target detection for anti-submarine warfare (ASW). It was based on information from various sensors, and was an attempt to systemize the art/craft of detecting enemy submarines. Information obtained from sensors includes distance, direction, angular velocity, sound patterns, doppler changes, and seabed topography. In the past, individuals determined the target's position based on experience and intuition.

However, as the number of sensors increased and their functions became more sophisticated, the information content became more complex and the amount of information significantly increased. Additionally, the high performance of submarines made it necessary to make decisions quickly. CLASSIFY processed a large amount of data sent by sensors, referred to the databased experience and intuition (knowledge base) of experts, and at the same time added statistical processing (fuzzy logic, Bayesian inference, etc.) to enable more precise target detection. The input information was not all processed sequentially, and a “blackboard” method was adopted where intermediate results from each part were written in a common area. This enabled parallel processing and shortened the overall analysis time. In other words, it was intended to overcome the volume problem (information volume, time constraints) in ASW information processing and improve quality (statistical processing).

ADRIES, developed as a monitoring system for ground-based targets, was expected to add additional information and statistical processing to the collected data, supporting tactical decisions (especially analysis of the nature and scale of enemy units). ADRIES combined ground search data collected from synthetic aperture radar (SAR) with terrain information as a pre-built knowledge base, basic unit composition, and other information obtained through reconnaissance, then statistically processed it to evaluate the situation and recognize targets. The combination of a knowledge base and statistical processing was the same as in CLASSIFY.

These systems use statistical processing to infer an overall picture from partial information by combining certain information with uncertain information. In other words, it systemized the work of “A word to the wise is enough (for appropriate guessing the whole).” In addition to deduction estimation, algorithms (Bayesian inference) became a major pillar in statistical information processing. Uncertain information is also processed with Bayesian inference, and it provides users (commanders, etc.) with results in the form of “There is an XX% probability that enemy tank unit is located in the highlands ahead, and the unit seem to be regiment size with the probability of YY%.” A similar system was also developed in Australia (Decision Support System).

3. Maintenance Support via Expert Systems

Apart from tactical decision support for combat commanders, systems incorporating expert systems were also considered for work decision support in logistics support divisions. The US Army developed PRIDE as a maintenance support system for failure recovery with the HAWK surface-to-air missile system's radar. The purpose of this system was to turn the know-how of veteran radar repair personnel into a database (knowledge base) so that even inexperienced soldiers could respond to and handle repairs to a certain extent when a failure occurred. PRIDE was brought to Saudi Arabia during the Gulf War (1990-91) and used for maintenance and repairs of locally deployed HAWKs. The US Army also developed a maintenance support system for M1 Abrams main battle tank.

The US Air Force built the EMMA missile maintenance support system using expert systems. Missile parts were modularized, so on-site repairs centered on module replacement. The system repeatedly did a question and answer routine with the technician to identify the replacement module for the failed missile. This was also work that required experience, and, with the example of GBU-15 guided bombs, using EMMA reduced time to identify repair points by 40% per bomb and by 74% overall. Cycle cost for the GBU-15 was also reduced.

According to the US Navy, the IDS maintenance support system for close-in weapon systems (CIWS), which used an expert system, reached a 91% success rate for locating failures. In addition, IDS discovered, in one hour, the cause of a failure that maintenance staff were unable to identify over the course of two weeks. Most of these maintenance support systems were designed to run on a personal computer, in consideration of convenience at work sites. In general, naval air force equipment tends to be multifunctional and limited in volume, but on the other hand, it is extremely difficult to identify the cause of failure by moving the equipment, spreading it out, and disassembling it, particularly when equipment on ships fails during operations at sea. It was thought that such maintenance obstacles could be overcome by using maintenance support systems that utilized AI (expert systems). In addition, improving on-site response capabilities makes it unnecessary to send equipment back every time a failure occurs, improving operational efficiency.

4. Issues with the 2nd AI Boom

For users, expert systems seemed to be consultants with excellent expertise, but the 2nd AI boom itself died down in the mid-1990s. There were two main reasons for this: problems specific to expert systems and problems related to the coexistence of humans and AI.

First, one of the problems specific to expert systems is that it is not easy to turn experts' knowledge into a database and make a "knowledge base." Experts' knowledge includes vague things like experience, intuition, and tacit knowledge, and there are many cases where experts answer by saying "We judge it this way in this case, but I can't adequately explain why." Programmatically expressing this requires a lot of effort, and the results are not necessarily what experts would expect. And for input information as well, experienced mechanics intuitively know that "the engine sounds different from normal," but when inputting this to an AI system, it is necessary to divide the sound into elements such as volume, frequency, period, directionality, sound quality, etc., and then define whether there is a statistically significant difference from normal cases.

In addition, for defense equipment, the same piece of equipment will be subject to frequent small-scale improvements as technology evolves and operational requirements change, and maintenance and repair methods will change accordingly. In other words, the maintenance support system also needs small scale improvements that match the equipment, and multiple maintenance support systems must be prepared if there are multiple versions of the same equipment. AI lacks the versatility that humans ordinarily have in this situation, so the system architecture is in danger of becoming complicated.

Additionally, it is not hard to imagine situations in which the information provided by the AI isn't useful for consulting. The aforementioned tactical decision support system provides information on theoretical probabilities, as can be seen from its use of Bayesian inference. However, information such as "The forward deployed enemy has a 60% probability of being a regiment, 30% probability of being a battalion, a 70% probability of being a tank unit, and a 25% probability of being a self-propelled artillery unit," even if it is provided to commanders, is not very helpful. Eventually a scout will be sent out and it will be directly confirmed.

Above all, there was the problem of the coexistence of humans and AI in intellectual labor. It is not uncommon for experts to have varying assessments of a particular situation and use different methodologies to make that assessment, yet the opinions of specific experts are reflected in the system. That being the case, the system does not become a reference for those who have differing opinions.

Reminiscent of the Luddite movement that took place in Britain in the early 19th century, technological innovation has always been confronted by concerns of "whether AI will deprive humans of intellectual roles." This is not just about working on behalf of AI or humans, but includes the fear of AI threatening human dignity in the form of "intellectual labor." This was the problem with MYCIN, which did in fact achieve good results. It was therefore not widely adopted in the medical field. Currently, we are entering the age of the 3rd AI boom, which is based on big data, deep learning, and inductive reasoning (see Table), and the matter of coexistence between AI and humans remains a major issue.

<References>

・Ingemar J. Cox et al., "Artificial-Intelligence Systems in Antisubmarine Warfare: Result of a Pilot Study with Expert Systems," *Saclantcen Memorandum* SM-176 (Dec., 1984).

・Allan M. Din ed., *Arms and Artificial Intelligence: Weapon and Arms Control Applications of Advanced Computing* (Oxford: Oxford University Press, 1987).

・Jude E. Franklin et al., "Expert System Technology for the Military: Selected Samples," *Proceedings of the IEEE* vol.76, no.10 (Oct., 1988).

・N. G. Gori and P. F. Calder, "A Decision Support System for Military Intelligence Assessment," *Technical Memorandum*, WSRL-TM-18/91 (Defense Science and Technology Organisation, Department of Defence, Australia: Mar., 1991).

・Robert J. Ivey Jr., "Diagnostic Expert Systems Use in the United States Navy," *Naval Postgraduate School Thesis* AD-A251 789 (Mar., 1992).

・Keishi Ono, "Replacement of Military's Intellectual Labor Using Artificial Intelligence: Discussion about AI and Human Co-existence," *NIDS Journal of Defense and Security*, no. 20 (Dec., 2019).

・Yutaka Matsuo, *Jinkochino wa ninge wo koeruka* [Will Artificial Intelligence Exceed Humans?], (Tokyo: Kadokawa, 2017).