Using Multi-Agent Systems to Pursue Autonomy with Automated Components

Jeffrey W. Tweedale\textsuperscript{a,b,*}

\textsuperscript{a}DSTO, PO Box 1500, Edinburgh, South Australia, 5111, Australia
\textsuperscript{b}KES, School of Engineering, ITEE, University of South Australia, Mawson Lakes, 5095, Australia

Abstract

Humans have used tools to transform raw resources into valued outputs ever since society harnessed fire. The type of tool, amount of effort and form of energy required depends on the output or object being created. As tools evolved into machines, they enhanced operator productivity. Hence, industry continues to invest heavily in machines to assist people to do more with less physical control and/or interaction. This involves automating functions previously completed manually. Taylorism and the Hawthorn experiments all contributed to optimising industrial outputs and value engineers continue to promote a mechanized workforce in order to minimise business variations in human performance and their behaviour. Researchers have also pursued this goal using Computational Intelligence (CI) techniques. This process of transforming cognitive functionality into machine actionable form has encompassed many careers. Machine Intelligence (MI) is becoming more aspirational, with Artificial Intelligence (AI) enabling the achievement of numerous goals. More recently, Multi-Agent Systems (MASs) have been employed to provide a flexible framework for research and development. These frameworks facilitate the development of component interoperability, with coordination and cooperation techniques needed to solve real-world problems. However problems typically manifest in complex, dynamic and often hostile environments. Based on the effort to seek or facilitate human-like decision making within machines, it is clear that further research is required. This paper discusses one possible avenue. It involves future research, aimed at achieving a cognitive sub-system for use on-board platforms. The framework is introduced by describing the human-machine relationship, followed by the theoretic background into cognitive architectures and a conceptual mechanism that could be used to implement a virtual mind. One which could be used to improve automation, achieve greater independence and enable more autonomous behaviour within control systems.

1. Introduction

Autonomy is a human term related to the rights and will of an individual. The definition currently refers to freedom of choice and there is no agreed definition when it is applied to machines. At present it is often used...
to describe automated behaviour. The linguistic definition of autonomy involves independence, freedom, self-governance and the rights or will of an individual [1], whereas, machines don’t have the ability to make conscious decisions, rationalise information or even recognise objects within a given context. They can only enforce pre-programmed decision-making and engage in limited forms of learning. Attempts have been made to provide machines with human-like decision-making capabilities, with limited success. Most research has been conducted within a constrained environment and focused on specific outcomes. This paper discusses one possible avenue of future research towards achieving a cognitive sub-system that may eventually reside and operate on an unmanned platform [2]. The concept of creating a virtual mind underpins efforts to improve automation by enabling greater independence and more autonomous behaviour within control systems.

A future with autonomous systems would introduce a ternary relationship between humans, machines and the level of effort or control required to achieve desired outcomes. A host of interrelated organisational and industrial functions need to be communicated. Multi-Agent Systems (MASs) are often used to provide components with sufficient capabilities required to solve complex issues. Such agent frameworks enable designers to abstract away system complexities, such as coordination, cooperation and communication [3]. Several of these capabilities rely on one or more Artificial Intelligence (AI) technologies and often involve hybridized solutions using agent teams.

Literature reviews reveal that the AI domain remains challenged with investigating and modelling complex real world problems when developing intelligent solutions. The field is being supported by ongoing increases in processing power, communication infrastructure and distributed paradigms. People increasingly use smart technologies in their business, social and professional lives. There is an increasing reliance on technology in how we live our lives and in hybrid solutions that holistically solve real world problems. For instance, technology can be used to process a sense, decide, act loop, although the decision making process is generally augmented through human intervention. Although modern sensors and actuators integrate processing abilities (making them capable of interacting with the environment), humans still provide most of the cognitive decision-making. Currently researchers rely on AI techniques to obtain information and help with decision making to solve isolated problems. Unfortunately these techniques currently do not exhibit sufficient intelligence to address complex cognitive problems. Intelligence implies elements of a humans’ ability to reason, learn and react based on ethics or values, including all traits that enable mammalian beings to adapt to new or unpredictable environments.

The AI domain also consists of many diverse sub-domains that form the current research frontier. Examples include: signal processing, knowledge representation (of structured and unstructured data), computational modelling, pattern recognition, data analytics, behavioural modelling and knowledge discovery which could be embedded in a virtual mind. These domains encompass research or innovation that aims to meet the emerging needs of industry and the community. The potential of AI is enormous and its full breadth is difficult to fully grasp, although it is believed that a golden age is beginning. This renewed interest in the modern era of AI is anticipated because of the domains’ evolving technical capability to solve complex real world problems through the use of MAS [3].

This paper provides a brief introduction into future research that is pursuing autonomy using automated components within a MAS framework. Section 2 provides background knowledge of the business models surrounding industry, while Section 3 discusses a number of concepts related to automation. Section 4 introduces issues surrounding Machine Intelligence (MI), the environment and its Governance. Section 5 discusses the concept of providing a cognitive architecture that may assist in delivering decision-making and possibly lead to independence or self-governance (autonomy). This discussion is followed by a description of the proposed conceptual framework in Section 6 and then a conclusion with comments about future research.

2. Background

As humans evolved, they slowly transitioned from a *hunter and gatherer* or *subsistence* society, to more structured communities that collectively provided for one another. Communities established trade-routes with other societies. Land owners formalised economic relationships to establish food assurance and domesticated plants and animals to promote production. Machines eventually improved economies of scale and their associated activities to stimulate the agricultural revolution. This led to the birth of the *producer-supplier* model of production (outputs).
As cottage industry gave way to organised production and eventually the industrial revolution, techniques promoted by Taylor [4] and Mayo [5] helped optimize workplace performance and motivation. Efficiency engineers used the Hawthorne effect to enhance production outputs within a supply model. Corporations also internationally championed this approach and thereby disrupted the supply model to generate outputs globally (disrupting disparate supply chains). Figure 1 illustrates the relationship between humans, machines and the outputs they produce.

![Diagram](image)

Fig. 1. The relationship between people, machines and their achievements or outputs

Organizationally, the original subsistence model grew vertically with segments relating to workers, management and outputs. Additional layers were introduced during the industrial revolution and, when communities started to populate cities, management controls were implanted to coordinate outputs in a new consumer/supply model. Slowly these segments then started to separate laterally under pressure from corporations who rationalised decentralized components into global models. The resulting consumer/provider/supply model is slowly transforming how society interacts with supply and its links to industry. Figure 2 depicts how machines are being used to increasingly wedge affected segments apart. Unfortunately this process is further isolating communication links and displacing people from existing production systems (outputs), which affects functional efficiencies because these systems were originally designed around people and not machines.

Value engineers continue to embrace technology to mechanise the workforce. This increased employment of machines continues to displace workers. The more primitive machines operate using closed loop systems; however, as complexity increases, additional systems or layers are wrapped around the central control loop to enable the controlled injection of external influences. A modern example could be a flight management system on an aircraft (fly-by-wire). Such systems typically maintain attitude control in real-time, although positional offsets can be injected to adjust the motion of the physical platform and its geographical location. Additional loops may support further automation, such as an autopilot function or augmented safety systems.

Humans increasingly rely on machines to assist them in manufacturing and support their everyday lives. During the post war era, industry embraced robotics and continues to mechanize its workforce. The trend began with logistic systems (packing, stacking, canning and bottling). Some mechanised functions are no-longer limited to static locations (painting, welding and assembly), creating a transition from machines being operated by humans to a paradigm where people simply supervise process flows. Recently machines are being controlled remotely through ’televisory’ links, while other systems have seamlessly automated functionality [6]. This has generated significant discussion over terminology. Currently people operate tools and supervise automated machines, however in the future humans may simply choose to employ an autonomous capability. One assumption is that machines are ’tools’ that transform energy into one or more actions [7]. Another is that machines evolved as a result of automation to enable more system capabilities [8]. This transformation is consistent with plans to improve manufacturing throughput and the quality of products, but displaces humans in the process. True autonomy
should be considered as self-governance, with a state or condition that enables independent decision making. This implies an evolutionary cycle that unifies automation and autonomy. They are not mutually exclusive terms, but operators will employ multiple systems (containing automated functionality) to enable more autonomous outputs.

The computer is an autonomous tool that can be used to extend the humans’ capacity to enhance productivity. Humans provide the cognitive stimulus as the primary decision maker, while the machine provides the processing capacity to automate functions previously completed manually. An example is productivity enabled by software tools like spreadsheets, word-processors and presentation applications. In production systems, MI automates processes and sequences requiring constant human intervention or control. Engineers are forced to anticipate the cognitive requirements for any derived task and synthesise suitable systems intended to provide human-like responses. If researchers want machines to engage in human-like activities, there is a need to synthesize the appropriate cognitive skills (possibly a primitive brain). This is a non-trivial task, with many complexities being tackled in isolation. Examples include: Defence Advanced Research Project Agency’s (DARPAs) Grand Challenge [9], their Systems of Neuromorphics Adaptive Plastic Scalable Electronics (SyNAPSE) program [10] and the Center of Excellence for Learning in Education, Science and Technology (CELEST) Modular Neural Exploring Travelling Agent (MoNETA) activity [11].

Based on the parallel nature of the brain and existing effort, more needs to be done to coordinate research in this domain. For instance making machines perceive information within their environment and then to recognise facts or knowledge. There is a need for machines to process environmental information and transform perceived knowledge within context to create awareness. Once information is recognised, categorised and transformed, a virtual brain should be able to engage in human-like reasoning against multiple criteria (ethics, values or even
beliefs) and then evaluate courses of action and rationalise about possible outcomes. If this process arrives at an acceptable solution, formal planning may begin before committing to a decision to act upon. Once the decision is communicated (internally or externally) the cycle of evaluation continues with possible refinement or adjustment based on the resultant environmental information (that accumulates over time), and to build a primitive level of awareness.

3. Existing Contributions

Humans currently perform a broad range of tasks that require a variety of skills. Each task may require the human to adopt one or more roles. It is possible that each role is fulfilled using a prescribed set of rules and the knowledge required provides focus and orients the mind-set required to achieve successful outcomes. The situation will determine the context by which decisions are made. Humans have typically been trained to cope with a select set of processes and procedures (complete with their own rules, boundaries and constraints). These skill sets will normally be employed routinely after acquiring a prescribed competency and level of expertise (now retained as knowledge). The ability to embed knowledge into a capability will determine the capacity of machines to fulfill cognitive roles. That employment is also reliant on societies’ acceptance or willingness to employ those machines. Prior to acceptance, researchers need to satisfy the requirement to provide machines with the ability to recognise and dynamically adapt to instil awareness of their environment.

General Motors installed the first industrial robot (Unimate) in its New Jersey factory in 1961. The mechanized work force continues to evolve. Today there are between 50–250 robots per 10,000 employees\(^1\) within industrial nations \cite{12}. According to the International Federation of Robotics\(^2\), this figure will grow by 200,000 installations per annum. After identifying propositions to increase scale and productivity, all BRIC countries (Brazil, Russia, India and China) have also declared plans to significantly increase future investment in robotic manufacturing.

Machines are still considered primitive with respect to human-level intelligence. Machines lack cognitive skills and this reduces industries’ ability to displace humans engaged in higher order activities. Unlike humans, there is no convenient method of comparing or categorising cognitive skills within machines. For instance, there isn’t a mechanism to measure machine intelligence or determine a Machine Quotient (MQ); however, it would be useful to be able to compare that score against an Intelligence Quotient (IQ) ranking. A benchmark of human-like intelligence may make the level of cognitive processing easier to identify and ultimately help derive a workable solution. There are many factors to consider before science can achieve self-governance or autonomy within machines. Such an achievement would produce a sound definition and result in real autonomy. At present, many industries still relying on automation to insulate them against labour shortages and demographic shifts. Many of the adopters are already reporting improvements in long-term sustainability, ecologically friendly production and power efficient measures that often leads to increased profit margins (globally). The author asserts it is possible for researchers to collaboratively employ “intelligent capabilities” using automated sub-systems to improve the level of autonomy. Increased investment is required to facilitate the use of more modern AI techniques and therefore promoting a shift towards improved automation. These improvements can be realised when humans use tools, machine assistants and even intelligent systems within the workplace.

As individuals participate in delivering specific activities, teams are often required to deliver larger outcomes. Many teams may be required to complete increasingly more complex tasks. When a team grows, organisations typically use predetermined hierarchies that contain embedded links to manage the group. An example includes activities; such as, the coordination of either, cooperation or Command and Control (C2). Additional mechanisms are then required to elicit efficient collaboration and situation awareness through cooperative processes.

The situation is generally monitored within the constraints of a mission or goal. There is a computational cost in maintaining or coordinating the number and variety of adjustments needed to enact or redirect the system goals. As with all complex systems, there is the potential to cause a ripple effect in which each action can create the possibility of unintended outcomes through emergent properties. These can trigger a need to review

\(^{1}\)In Australian terms this equates to only 0.5%–2.5% of its total workforce.

\(^{2}\)These statistics are reported at www.worldrobotics.org
outstanding plans and their associated actions to better reflect the equivalent human processes. This declarative knowledge also needs to be assimilated by engineers within a final product offering. An inference system can be incorporated to assist machines to exhibit human-like behaviors. There are a number of existing models, based on concepts like the Joint Director of Laboratories (JDL) data fusion [13], Boyd’s Observe, Orient, Decide and Act (OODA) loop and Beliefs, Desires and Intent (BDI) agents [14]. These models and techniques should be reviewed to reflect a dynamic environment based on complex scenarios in the real-world. For instance, machines sense the environment, where humans observe their surroundings. To achieve human-like computation, the machines’ primitive mind (using an internal loop to be used for cognitive thought) might include:

- Sense (Observe)
- Detect (Perceive)
- Acquire (Recognise)
- Categorise (Rationalise)
- Orient (Contextualise)
- Transform (Reason)
- Choose (Determine)
- Commit (Act)

Boyd’s OODA loop describes the human cognitive processes (in terms of wet-ware or Gray-ware) whereas machines use BDI agents to incorporate decision-making processes. This is limited to a sense-decide-act loop within the computational components. Machines only achieve tasks they have been programmed for and commanded to achieve certain goals. It is therefore difficult for machines to cope in unknown environments, especially when humans provide off-board decision making. Ultimately, elements of the BDI loop must be augmented using the functionality associated with the OODA loop. It is not clear what the final system will be called and/or what it will contain (See Section 4), however there is evidence that a virtual mind will potentially enable more autonomous decision making on-board unmanned platforms (See Section 5).

Information only becomes knowledge after data is collected, integrated and processed. Raw data needs to be perceived and recognised prior to being rationalised against a given context. If the system is aware of new information or facts, it can adapt the process based on extant knowledge or existing rules before consciously assessing the ramifications. This may provide a rationalising effect by considering a series of possible outcomes before committing to a specific decision or action. Humans traditionally gather information from their environment and process it internally before committing to a specific course of action. Consistent with this, researchers seek to embed cognitive functions within machines, but humans are often still required to mediate and communicate decisions before machines can act. Figure 3 depicts the interactions for a complex and dynamic environment within a real-world context.

![Fig. 3. Real-world Interaction](image)
Human-computer systems are sometimes modelled as continuous loops that sense environmental information. That information may be processed and accumulated in a knowledge-base or repository. Other functions may be applied to interpret or rationalise new facts against embedded or acquired rules. The rules need to be adapted to operate within the context of a situation and based on the context or tempo of each mission or goal.

4. The Machine Intelligence Environment

The origins of MI appear to have been consolidated at a workshop at Carnegie Mellon University in 1980. Given the overlapping effort of research using both AI and cognitive science to elicit learning, a journal was conceived to focus on methods used to improve machine performance of a nominated task(s). This includes the representation, acquisition, structuring, categorization and classification of data to extrude knowledge [15]. It now includes techniques that sponsor symbolic representation and machine vision. Early topics included: production rules, decision trees and procedural reasoning. The original definition of MI was expanded to improving performance through experience, for instance using techniques like pattern recognition and data-mining or knowledge refinement (through probabilistic instance-based representation). More recently there is growing emphasis on knowledge classification and inductive logic (for regression tasks). In a recent editorial, Langley suggests it is time for Machine Learning (ML) to focus on “developing intelligent systems that exhibit the rich behaviour of complex tasks and acquire knowledge cast in rich relational structures” [16]. Many scientists consider that on-board MI is essential for efficient recognition and understanding within complex or dynamic environments [17]. Some agree that further research into augmented reality [18] is one possible avenue of generating rich meta data that is often used for machines to learn or adapt within a given context. Another potential push is to exploit contextual databases like the ‘Cycorp’ semantic web [19]. An alternative is to pursue research into cognitive architectures or hybrid reasoning systems.

The DARPA recently commissioned several projects that influence AI developments, including a series of autonomous car demonstrations. For instance, in 2006 a team of researchers employed a Volkswagen Touareg to successfully cover 212 kilometres across the Mojave desert to win the Grand Challenge. This autonomous car utilized a complex array of brain-inspired microprocessors embedded into fourteen blade severs. The competition was extended in 2007, and all challengers (their machines) failed the common knowledge test. This test called for machines to demonstrate basic driving tasks, such as merging, passing, parking, negotiating intersections and avoiding random objects (such as rocks or pedestrians). It is clear that the sensors fitted to the machines enabled them to perceive changes within its environment, but their decision-making system(s) failed to recognise random threats introduced throughout the challenge. This example indicates that despite the processing capacity of each machine, their design(s) continue to fail primitive survival challenges. The same challenges a rat with only two grams of grey matter achieves routinely.

Researchers endeavour to replicate the brain using silicon systems. These machines use significant Central Processing Unit (CPU) resources, and few have managed to emulate the brain (The MoNETA Project synthesized a small mammal in 2010. This contained over 60,000 neurons and 120 million synaptic connections [11]). A human brain has 100 billion neurons and, regardless of cost or resources, an emulated equivalent would consume the power of a medium sized town to switch it on. This research acknowledges that the human brain runs as a massively parallel system and conducts knowledge processing in-line with any associated data. Even with further research in the MI domain, cognitive architectures will rely on improvements in Computational Intelligence (CI) techniques. Following these achievements, it may be possible to combine the scientific understanding of the brain with improved parallel processing techniques in software in order to further investigate a virtual application that enables machines to make higher-order cognitive decisions.

5. Cognitive Architectures

A cognitive architecture provides a framework upon which physical symbolic systems may be realised. It provides definitions, resources, constraints and management processes needed to achieve cognitive processing.

---

3See the Machine Learning Journal, originally published by Kluwer, however the editors have since championed the Machine Learning Journal - Research at http://jmlr.csail.mit.edu/ (which is free on-line)
A number of cognitive architectures have evolved, including Procedural Reasoning System (PRS) [20], Soar [21], Adaptive Character of Thought – Rational (ACTR) [22], ICARUS [23], PRODIGY [24], distributed Multi-Agent Reasoning System (dMars) [25] and a suite of BDI [26] agent frameworks. Each seeks to develop a form of perception that learns environmental information based on human-like or a cognitively focused architecture (mimicking human-like behaviours) to solve realistic problems.

Some architectures failed because they attempted to map and reproduce (or personify) human-like functions and procedures. It may be possible to stimulate machines to provide appropriate responses without the need to duplicate or mimic human cognition, although recognition, learning, memory (persistent and temporal), reasoning and rationality are key attributes that must be pursued. Figure 4 shows the taxonomy for a simplified cognitive architecture proposed by Dutch et al. in 2008 [27].

![Fig. 4. Simplified Taxonomy for a Cognitive Architecture [27]](image)

A group of researchers continue to expand the ICARUS cognitive architecture. This software was originally developed at the Stanford University in the 1990’s. It is an agent framework that allows researchers to host a system with human-like decision making capability. ICARUS specifies the learning behaviour of AI agents, where new skills are identified following on from a success when problem solving [28]. Its architecture is segmented into modular components: a perceptual system (ARGUS), a planning system (DAEDALUS), an execution system (MAEANDER), and the memory system (LABYRINTH). It uses a Markov Decision Process (MDP) model to map convergent activities based on the State Hierarchy, Action, Reward, State Hierarchy, Action (SHARSHA) re-enforcement algorithm [29]. It stores the short-term situation (accumulated status) and long-term knowledge (learned behaviour) in buffers as a skills hierarchy. Based on the literature, few of the existing architectures have used or considered more modern constructs; such as coordination, cooperation or the notion of teams within organisational hierarchies.

6. Conceptual Framework

As suggested above, this paper is aimed at creating discussion about the research required to implement a cognitive framework. Figure 5 represents the authors conceptual design of an agent framework that might be used to support future research into human-like aspects of decision making. This framework could employ hybrid cognitive architectures that can support a virtual mind.
The MAS system is the central feature of the framework. Each agent within the system will encapsulate one or more capabilities, while hidden elements abstract away the enterprise level functionality. Layers of the agent system provide interoperable functions, such as logging, communications, migration and any other interfaces required by the parent application. The agents operate within a centralized (services-based) knowledge repository, that is structured by category and allows the agent system to seamlessly perform context switching, while maintaining a real-time interpretation of the world. As with all knowledge systems, data must be transformed into information. That information must be categorised and validated as relevant knowledge that is accessible by the system. The author believes an inference engine should also be integrated within the MAS to assist in the promotion and interrogation of the growing knowledge repository. The agent system interfaces with both the operator (traditionally humans) and other clients. Future systems could aim to provide autonomous operators, which could also be clients of other systems. The tools at the bottom of the figure may vary, according to the application, and those provided as examples.

7. Conclusion and Future

Modern concepts related to CI and AI may eventually support human-like MI. The development of a virtual mind is non-trivial and significant effort is still required to achieve cognitive decision making on-board unmanned
systems. Analysis of existing cognitive architectures is ongoing. Once complete, modern technology can then be incorporated to exploit hybrid techniques in the agent system environment.

More effort is required to elicit a feasible structure (possible future standard) that represents both information and knowledge obtained within a dynamic, real-word environment. This concept relies on the MAS framework to abstract away the enterprise level complexity. The latter is aimed at hiding the system complexity and enable less qualified researchers to streamline the employment of new capabilities. The author believes that once the MAS framework is fully investigated, a cognitive architecture would evolve. Experiments can be used to validate the final capability and facilitate the transformation of declarative knowledge within a virtual mind.

References