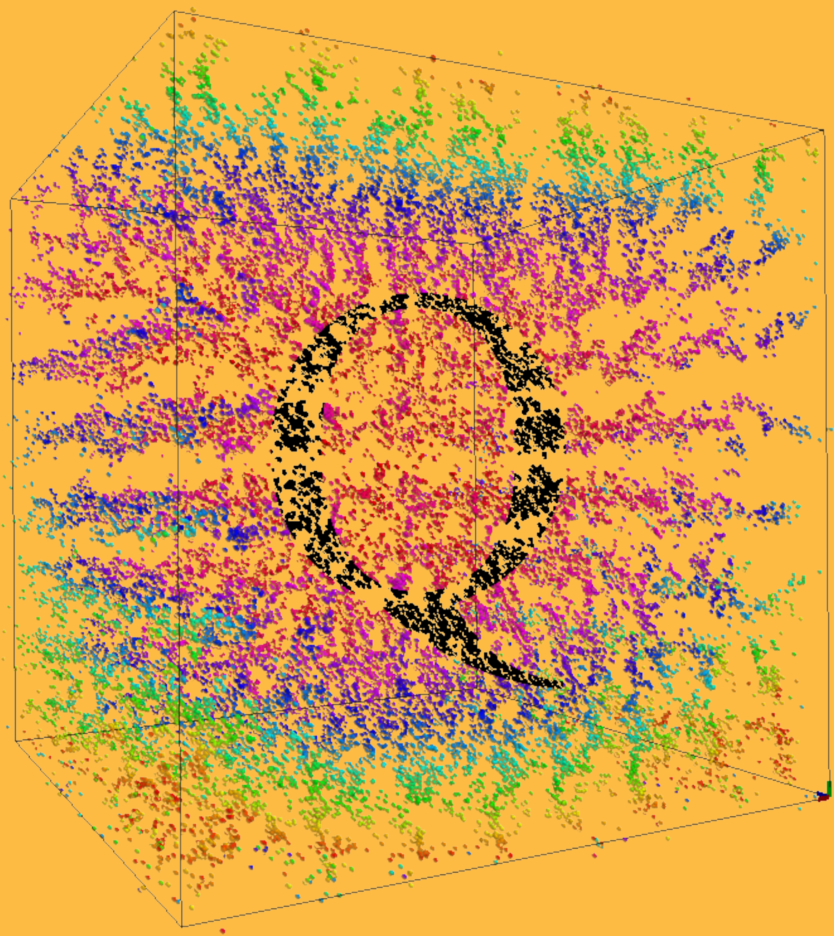
AISB QUARTERLY

THE NEWSLETTER OF THE SOCIETY FOR THE STUDY OF
ARTIFICIAL INTELLIGENCE AND SIMULATION OF BEHAVIOUR



No. 141 July, 2015

Do you feel artistic?
Exhibit your artwork on our front covers!
Email us at aisbq15@aisb.org.uk!

Artwork by Alwyn Husselmann, PhD (Massey Univ., New Zealand)

Visualisation is an important tool for gaining insight into how algorithms behave. There have been many techniques developed, including some to visualise 3D voxel sets [1], program space in genetic programming [2] and vector fields [3], amongst a large number of methods in other domains. A qualitative understanding of an algorithm is useful not only in diagnosing implementations, but also in improving performance.

In parametric optimisation, algorithms such as the Firefly Algorithm [4] are quite simple to visualise provided they are being used on a problem with less than four dimensions. Search algorithms in this class are known as metaheuristics, as they have an ability to optimise unknown functions of an arbitrary number of variables without gradient information. Observations of particle movement are particularly useful for calibrating the internal parameters of the algorithm.

Pictured on the cover is a visualisation of a Firefly Algorithm optimising the three-dimensional Rosenbrock Function [5]. Each coloured sphere represents a potential minimum candidate. The optimum is near the centre of the cube, at coordinate $(1, 1, 1)$. Colour is used here to indicate the output of the Rosenbrock function, whereas the 3D coordinate of each particle is representative of the actual values used as input to the function. The clustering seen amongst the particles in the image is due to the local neighbourhood searches occurring. Particles tend towards the optimal solution in a small group of a certain radius, as well as moving randomly to a certain degree. The visualisation assisted in improving convergence and verifying the implementation of the optimiser.

[1] KA Hawick, '3d visualisation of simulation model voxel hyperbricks and the cubes program', Technical Report CSTN-082 102-904, Massey University, Albany, North Shore, Computer Science, Massey University, Albany, North Shore 102-904, (October 2010).

[2] AV Husselmann and KA Hawick, '3d vector-field data processing and visualisation on graphical processing units', in Proc. Int. Conf. Signal and Image Processing (SIP 2012), pp. 92Ð98, Honolulu, USA, (August 2012).

[3] AV Husselmann and KA Hawick, 'Parallel parametric optimisation with firefly algorithms on graphical processing units', in Int. Conf. on Genetic and Evolutionary Methods (GEM'12), pp. 77–83, (2012).

[4] AV Husselmann and KA Hawick, 'Visualisation of combinatorial program space and related metrics', in 2013 International Conference on Information and Knowledge Engineering (IKE'13), (2013).

[5] XS Yang, 'Firefly algorithms for multimodal optimization. in: Stochastic algorithms: Foundations and applications', in SAGA, (2009).

Editorial

To quote an anonymous Committee member, as spotted on one of the social websites that dominate our lives:

> I'm back! Did you miss me? What do you mean 'you didn't notice I was gone'? Sure you did. Sure you did.

The truth is, the world is changing at a pace faster than we are able to cope. Gigabit ethernet is now (almost) the internet of the past. It will soon be the solution that my three year old son will have to fall back on, when his head-mounted, miniaturised computer is not able to download the world news, on his daily commute between London and New York[1].

At the time I am writing these lines, Universities are trying to adapt to this evolving market, as it is now called. Departments are closing faster than new ones are being open. National evaluation exercises and league tables dominate the academic life and quite literally redefine what it means to pursue academic fulfilment, which is now a privilege of a few.

On the plus side, AI is booming again, thanks to both progress made by industry, and great outreach threads, including amazing PR from companies and a series of big bucks movies. I guess this is a good thing for our field more broadly, if we are in a position to leverage this hype.

Hopefully, you have noticed some delays in receiving your Quarterly. The truth is, it has always been difficult to recruit people to submit material to be pushed to our members, but these days it's becoming increasingly difficult. We are competing for time, and it's a battle we will not win.

AISB has about 400 members, mostly in the UK. In times like these, the Quarterly should thus be seen as an outreach tool, designed to bring you followers and "Likes"; something tangible, palpable that we are naturally inclined to fall back on. I can only invite you to get in touch with us, and to make your voice heard.

∴

Echoing my editorial, Joel Lehman et al. review the debates that happened on the occasion of a recent AAAI workshop, which gathered leading figures. Their conclusion is without appeal, the *anarchy of methods* that makes AI is to be embraced. Significant and important challenges remain, however, including safety, ethical and societal considerations.

Andy Thomason reviews the recent book by van Benthem on logic in games, and JeeHang Lee and Jekaterina Novikova report on conferences they attended to, thanks to funding from AISB.

Etienne B. Roesch
Editor-in-Chief

[1] I wonder if Fireman Sam will be viewable with 3D glasses then...

An Anarchy of Methods: Current Trends in How Intelligence is Abstracted in AI

by *John Lehman* (U. Texas, USA), *Jeff Clune* (U. Wyoming, USA), &

Sebastian Risi (U. Copenhagen, Denmark)

Artificial intelligence (AI) is a sprawling field encompassing a diversity of approaches to machine intelligence and disparate perspectives on how intelligence should be viewed. Because researchers often engage only within their own specialized area of AI, there are many interesting broad questions about AI as a whole that often go unanswered. How should intelligence be abstracted in AI research? Which subfields, techniques, and abstractions are most promising? Why do researchers bet their careers on the particular abstractions and techniques of their chosen subfield of AI? Should AI research be "bio-inspired" and remain faithful to the process that produced intelligence (evolution) or the biological substrate that enables it (networks of neurons)? Discussing these big-picture questions motivated us to organize an AAAI Fall Symposium, which gathered participants across AI subfields to present and debate their views. This article distills the resulting insights.

Introduction

While researchers in AI all strive to create intelligent machines, separate AI communities view intelligence in strikingly different ways. Some abstract intelligence through the lens of connectionist neural networks, while others use mathematical models of decision processes or view intelligence as symbol manipulation. Similarly, researchers focus on different processes for generating intelligence, such as learning through reinforcement, natural evolution, logical inference, and statistics. The result is a panoply of approaches and subfields.

Because of independent vocabularies, internalized assumptions, and separate meetings, AI sub-communities can become increasingly insulated from one another even as they pursue the same ultimate goal. Further deepening the separation, researchers may view other approaches only in caricature, unintentionally simplifying the motivations and research of other researchers. Such isolation can frustrate timely dissemination of useful insights, leading to wasted effort and unnecessary rediscovery.

To address such dangers, we organized an AAAI Fall symposium that gathered experts with diverse perspectives on biological and synthetic intelligence. The hope was that such a meeting might lead to a productive examination of the value and promise of different approaches, and perhaps even inspire syntheses that cross traditional boundaries. However, organizing a cross-disciplinary symposium has risks as well. Discussion could have focused narrowly on intractable disagreements, or on which singular abstraction is "the best." An unhelpful slugfest of ideas could have emerged instead of collaborative cross-pollination, leading to a veritable AI Tower of Babel.

In the end, there were world-class keynote speakers spanning AI and biology (see below), and participants were indeed collaborative. Some travelled to the United States from as far as Brazil, Australia, and Singapore; but beyond geographic diversity, there were representatives from many disciplines and approaches to AI (Figure 1). Drawing from the symposium's talks and events, we now summarize recent progress across AI fields, as well as the key ideas, debates, and challenges identified by the attendees.

- Andrew Ng (Stanford U.): Deep Learning
- Risto Miikkulainen (U. Texas): Evolving Neural Networks
- Pierre-Yves Oudeyer (Inria, France): Developmental Robotics
- Gary Marcus (NYU): Cognitive Science
- Georg Striedter (UC Irvine): Neuroscience
- Randall O'Reilly (U. Colorado): Computational Neuroscience

Key ideas discussed

One controversial topic was deep learning, which has recently shattered many performance records over an impressive spectrum of machine learning tasks [2, 7]. The central idea behind deep learning is that large hierarchical artificial neural networks (ANNs), inspired by those found in the neocortex, can be trained on big data (e.g. millions of images) to learn a hierarchy of increasingly abstract features (Figure 2; [4]) Overall, participants agreed that recent progress in deep networks was a significant step forward for processing streams of high-dimensional raw data into meaningful abstract representations, e.g. recognizing faces from unprocessed pixel data. But there was also agreement that much work yet remains to create algorithms that leverage such representations to produce intelligent behavior and learn in real-time from feedback; in other words, scaling deep learning to more cognitive behavior may prove problematic.

Andrew Ng, affiliated with Stanford University and Baidu Research, gave a keynote on deep learning that outlined its motivation, implementation, and recent successes. Other keynote speakers reported that they also effectively use deep learning, in that their research similarly involves learning in many-layered neural networks. In this sense, deep learning has gone by many names over time, and is currently being reinvigorated by increased computing power, big data, greater biological understanding, and algorithmic advances. For example, in his keynote, Randall O'Reilly of the University of Colorado at Boulder summarized his work in the field of computational neuroscience, where researchers often develop cognitive architectures, which are computational processes designed to model human or animal intelligence. His Leabra cognitive architecture is a many-layered neural network modeled on the human brain, which includes collections of neurons analogous to the major known functional areas of the brain [10]. In

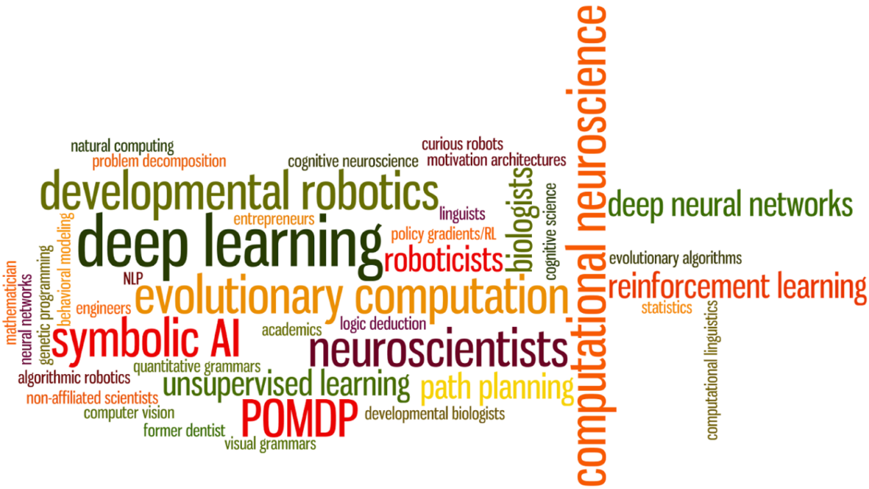


Figure 1: The backgrounds of attendees.

this way, two separate areas of AI apply similar technologies inspired by different motivations: one coarsely abstracts brains to solve practical problems, and the other applies more biologically plausible abstractions to better understand animal brains.

A related camp (to which the authors belong) that is inspired by nature and applies evolutionary algorithms to design neural networks, is called neuroevolution. In his keynote, Risto Miikkulainen of the University of Texas at Austin described how neuroevolution can design cognitive architectures via a bottom-up design process guided by evolutionary algorithms instead of through top-down human engineering. Kenneth Stanley, from the University of Central Florida, argued that evolutionary approaches may be important tools for producing human-level AI because evolution is highly adept at creating variations on an underlying theme [14]). The idea is that evolutionary methods could perhaps provide this important capability to other AI techniques, such as deep learning. Supporting this idea, Jeff Clune, from the University of Wyoming, described how evolutionary algorithms that incorporate realistic constraints on natural evolution can produce ANNs that have important properties of complex biological brains, like regularity, modularity, and hierarchy [3].

Pierre-Yves Oudeyer of Inria detailed in his keynote the field of developmental robotics, which investigates how robots can develop their behaviors over time through interacting with the world, just as animals and humans do [8]. Representative approaches in developmental robotics implement

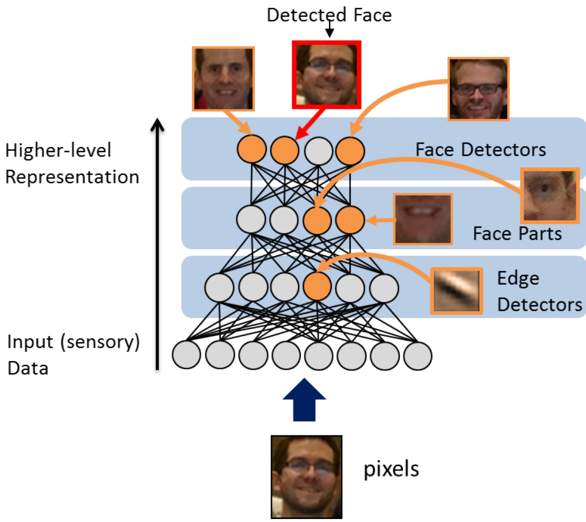


Figure 2: An illustration of deep learning. As a deep network of neurons is trained to recognize different faces, the neurons on the lowest level learn to detect low-level features such as edges, and higher-level neurons combine these lower-level features to recognize eyes, noses, and mouths. Neurons on the top of the hierarchy can then combine such features together to recognize different faces.

mechanisms to enable lifelong, active, and incremental acquisition of both skills and models of the environment, through self-exploration or social guidance. Oudeyer's research shows that motivating robots to be curious results in continual experimentation: A robot equipped with intrinsic motivation will search for information gain for its own sake; at any given point in the robot's development, it actively performs experiments to learn how its actions affect the environment [11]. Because such curiosity leads to an ever-improving model of the consequences of a robot's actions, over time it can result in learning how to accomplish increasingly complex tasks.

Among the traditional biologists that attended was Georg Striedter of the University of California at Irvine, author of the influential book "Principles of Brain Evolution". His keynote focused on the history of how brain functionality has been viewed over time. He noted an interesting parallel between the history of AI and of neuroscience: In both, a simple serial view of intelligence led to exploring more parallel, distributed notions of processing. He mentioned that Rodney Brooks' subsumption architecture in particular had resonated with him, because it offered a picture of higher-order thought beyond simplistic linear pathways (Brooks 1986); while computer scientists often debate the promise of various approaches to computational intelligence amongst themselves, it is informative also to consider the opinions of those who study how it arose in humans.

Aside from models with concrete biological inspiration, other attendees focused on abstractions of intelligence based on Markov decision processes (MDPs) and less-restrictive generalizations called partially observable Markov decision processes (POMDPs). Such MDPs and POMDPs represent decision making in a mathematical framework composed of mappings between states, actions, and rewards. This framework provides the basis for AI techniques like reinforcement learning and probabilistic graphical models. Devin Grady of Rice University and Shiqi Zhang of Texas Tech University each described mechanisms to augment such techniques to allow them to better scale to more complex problems. A similar need for tractable models motivated Andrew Ng's change in focus from MDP-based reinforcement learning to deep learning. He mentioned in response to a question that he felt the bottleneck was no longer reinforcement learning algorithms themselves, but in generating strong relevant features from raw input for such algorithms to learn from, which otherwise must be manually generated by humans through domain-relevant knowledge.

Proponents of symbolic AI (also known as GOFAI, or "good old-fashioned artificial intelligence," due to its early research dominance) defended their view that the power of human intelligence is largely captured in the idea of symbol manipulation. Such researchers also illuminated where non-symbolic approaches still fall short. In particular, John Laird of the University of Michigan posed an interesting challenge problem called embodied taskability: Similar to learning from demonstration [1], a robot must learn to perform novel tasks by interacting with humans. The task is intriguing because it is an ambitious problem not often tackled by other fields of AI, yet is characteristic of human intelligence. Complementarily, Gary Marcus of New York University gave a provocative keynote highlighting several capabilities necessary for strong AI that current high-performing connectionist approaches do not yet implement, such as representing causal relationships and abstract ideas, and making logical inferences. He also mentioned challenges in natural language understanding.

During one of the panel sessions, an idea was proposed in an attempt to tie all of these fields and levels of abstraction together: a stack of models, where each individual level of the stack is guided by a different level of abstraction. The idea is that with such a stack the various levels of abstraction could be linked together, guided by a reductionist goal of connecting understanding of high-level, abstract, rational components of intelligence to "lower-level" ones that are closer to perceiving raw data and controlling muscles. For example, high-level GOFAI algorithms could possibly be connected to deep learning models, which could be connected to more biologically-plausible computational models of brains. In this way it might be possible to unite disparate views and approaches to gain greater overall understanding.

Debates

As mentioned previously, deep learning proved to be a lightning rod for discussion and many researchers were quick to point out perceived difficulties in scaling deep learning to human-level AI. Open research questions include how to create deep networks that implement reinforcement learning, develop higher cognitive abilities over time, or manipulate symbols. Andrew Ng, when asked about merging deep learning with reinforcement learning, responded that it is an unsolved problem and that "a seminal paper on that subject is waiting to be written1." Ng was hopeful that it should be possible to extend deep learning algorithms to perform reinforcement learning without merging in other AI paradigms. In contrast, and perhaps unsurprisingly, researchers outside of deep learning were generally more skeptical.

While the current winds of AI seem generally to favor statistical machine learning methods like deep learning or reinforcement learning over purely symbolic GOFAI approaches, proponents of symbolic AI made convincing arguments for its continued relevance. John Laird expressed that although symbolic AI might not be as dominant as it once was, research progresses onward irrespective of current fashion. In particular, symbolic AI research is currently producing promising symbolic cognitive architectures that can empower agents to learn new human-taught tasks. In his keynote, Gary Marcus argued that it would be a mistake to conflate the time of an approach's first prominence with its potential; he noted that symbolic AI techniques might also (like statistical techniques) benefit from advances in computing power and available data, and that such symbolic techniques were developed mainly in the absence of the broad computational resources that are now used in statistical approaches.

A point of agreement was that symbolic AI is not better or worse that alternate approaches, but is instead different in its aims and objectives. Symbolic AI continues to aim at the ambitious goal of general artificial intelligence (i.e. human-level intelligence) while other approaches often focus on narrower domains or simpler forms of intelligence. A contribution of Gary Marcus was to highlight that GOFAI is not an inferior way of reproducing these narrower or simpler intelligences, but is instead aimed at a different goal: the cognitive intelligence that sets humans apart from other animals, which is where statistical machine learning methods are arguably weakest.

A contentious issue for researchers in biologically-inspired AI concerned which biological details are extraneous and therefore unnecessary to include in AI models. For example, brains vary over a multitude of dimensions including neuron size, density, type, connectivity, and structure; intuitively, it seems unlikely that all such dimensions are equally important to a model's functionality. Randall O'Reilly mentioned that in his models, the additional complexity of simulating neurons with binary spikes in time (like biological neurons) provided little benefit over simpler neuron models. Yet in contrast, Oliver Coleman (University of New South Wales) highlighted past research showing that the timing of spikes may be an important facet of learning processes in the brain. Taking the attendees as a whole, there were more skeptics of complex neuron models than proponents, likely reflecting a cautious, pragmatic preference for simplicity over biological realism for its own sake.

The opposite question was also debated: Are there salient features of brains and intelligence that are unfairly ignored? For example, O'Reilly believes that glial cells, which are non-neural cells that provide support and protection for neurons, may be more important computationally than their absence in most models would suggest. For Risto Miikkulainen and Pierre-Yves Oudeyer, how brains physically develop over time was a topic deserving greater attention; most models ignore the fact that biological brains learn while they grow and develop into their full mature size. In contrast, Gary Marcus argued that it may be possible to abstract nearly all biological detail away if all we care about is engineering AI, and not understanding biology. The resulting discussion questioned whether the brain is a well-engineered machine with much to teach us, or whether it is merely a hacked-together "kluge" [9]. In other words, do researchers mistakenly idealize the human brain, searching for elegant insights in a messily-designed artifact, one that is functional but ultimately unintelligible?

As the debate became more intense, Pierre-Yves Oudeyer interjected that, of course, which biological details are important depends upon the scientific question being investigated. Or, as John Laird said in response to the name of the symposium ("How Should Intelligence be Abstracted in AI Research?), "It depends!" Oudeyer then said something that resonated strongly: Because we do not deeply understand intelligence or know how to produce general AI, rather than cutting off any avenues of exploration, to truly make progress we should embrace AI's "anarchy of methods."

Major Challenges

Through the course of the discussion, many remaining challenges for AI became evident that cut across traditional boundaries. Overall, AI approaches tend to have four distinct focuses: Real-world embodiment, building features from raw perception, making decisions based on features, and high-level cognitive reasoning that is unique to humans. Approaches generally specialize on one such area, and often perform poorly when stretched beyond that focus. However, general AI requires spanning such divides. To do so may require integrating existing disparate technologies together; for example, hybrid neural systems [15] often combine neural network and symbolic models together, like the SAL architecture that connects the symbolic ACT-R model to bottom-up perception from the Leabra neural model [5]. A more conventional approach is to attempt to scale up an existing technology beyond its current borders. For example, Risto Miikkulainen's keynote highlighted that neuroevolution techniques are beginning to evolve instances of simple cognitive architectures. Additionally, cognitive architectures like Leabra and Spaun are beginning to tackle symbolic manipulation of variables through human-engineered neural mechanisms [13, 6]. Extensions to deep learning might similarly incorporate decision making and cognition. However, if integrating or extending existing technologies proves unproductive, there might yet be a need for new approaches better able to bridge aspects of AI ranging from low-level perception to human-level cognition.

An interesting challenge in AI that often goes unconsidered is safety. The most interesting intellectual challenge drawing researchers to AI is understanding and engineering intelligent systems. However, it may be dangerous to single-mindedly pursue such a goal without considering the transformative consequences that may result if we create AI that rivals or even surpasses human intelligence. Problematically, academic and industrial incentives are nearly unilaterally aligned towards creating increasingly sophisticated AI and do not emphasize critical reflection regarding its potential downsides or side-effects. Only a single talk, by Armando Tacchella of the University of Genova, focused on creating safe abstractions of AI [12]. That work raised difficult questions for the many AI approaches where verification or automatic characterization of the behaviors produced is difficult. For example, neural networks are notorious for being black box models, making interpreting the safety of agents resulting from deep learning, neuroevolution, and neural-based cognitive architectures difficult. A consensus among attendees was that this was an important and underfunded consideration.

Another central problem that emerged through discussions is the difficulty (or impossibility) of definitively knowing what ways of abstracting intelligence are truly "better" or more productive than others. In general, attempting to predict the future promise of any particular technology or research direction is often misleading. But a particular challenge in AI stems from the existence of only one example of high-level intelligence from which to infer generalities. As a result of nature's singular anecdote on intelligence, separating what is essential for intelligence from what is merely coincidental remains difficult.

Conclusion

At the symposium's end, researchers mentioned that they better understood the philosophical and theoretical motivations for areas of AI they had unintentionally only seen previously in caricature. One participant said that he learned that even when viewing intelligence abstractly from a high level, there is a benefit to following key developments at lower levels. Another offered that he "learned how limited our knowledge is," and that it was interesting how often "key leaders in a field might not have a grand, deep plan [...] but that instead, behind the curtain, are scientists doing the best they can, fumbling in the dark." Another made reference to the parable of three blind men describing an elephant, where each blind man describes the whole elephant in terms of features specific to the individual parts they are examining (a tail, a tusk, or a leg, respectively), which leads to very different interpretations of what an elephant is. Similarly, through sharing local perspectives on AI and what they imply about the overall field, the resulting traces of intelligence's outline – made from all angles and levels of abstraction of AI's anarchy of methods – might potentially be combined to accelerate our understanding of the general principles underlying intelligence and how to recreate it computationally.

Reference

[1] CG Atkeson and S Schaal, 'Robot learning from demonstration', in *ICML*, volume 97, pp. 12–20, (July 1997).

[2] Y Bengio, 'Learning deep architectures for ai', *Foundations and trends in Machine Learning*, **2**(1), 1–127, (2009).

[3] J Clune, J-B Mouret, and H Lipson, 'The evolutionary origins of modularity', *Proceedings of the Royal Society B*, **280**(20122863), (2013).

[4] GE Hinton, 'Learning multiple layers of representation', *Trends in cognitive sciences*, **11**(10), 428–434, (2007).

[5] DJ Jilk, C Lebiere, RC O'Reilly, and JR Anderson, 'SAL: An explicitly pluralistic cognitive architecture', *Journal of Experimental and Theoretical Artificial Intelligence*, **20**(197-218), (2008).

[6] T Kriete, DC Noelle, JD Cohen, and RC O'Reilly, 'Indirection and symbol-like processing in the prefrontal cortex and basal ganglia', *Proceedings of the National Academy of Sciences*, **41**(16390-16395), (110).

[7] A Krizhevsky, I Sutskever, and GE Hinton, 'Imagenet classification with deep convolutional neural networks', in *Advances in neural information processing systems*, pp. 1097–1105, (2012).

[8] M Lungarella, G Metta, R Pfeifer, and G Sandini, 'Developmental robotics: a survey', *Connection Science*, **15**(4), 151–190, (2003).

[9] Gary Marcus, *Kluge: The Haphazard Evolution of the Human Mind*, Houghton Mifflin Harcourt, 2008.

[10] RC O'Reilly and Y Munakata, *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*, MIT press, 2000.

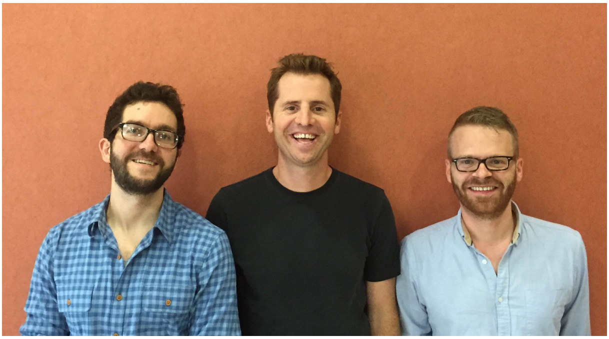
[11] PY Oudeyer, F Kaplan, and VV Hafner, 'Intrinsic motivation systems for autonomous mental development', *IEEE Transactions on Evolutionary Computation*, **11**(2), 265–286, (2007).

[12] S Pathak, L Pulina, G Metta, and A Tacchella, 'How to abstract intelligence? (if verification is in order)', in *Proceedings of the 2013 AAAI Fall Symposium on How Should Intelligence be Abstracted in AI Research*, (2013).

[13] NP Rougier, D Noelle, TS Braver, JD Cohen, and RC O'Reilly, 'Prefrontal cortex and the flexibility of cognitive control: Rules without symbols', *Proceedings of the National Academy of Sciences*, **102**, 7338–7343, (2005).

[14] KO Stanley, 'Compositional pattern producing networks: A novel abstraction of development', *Genetic Programming and Evolvable Machines*, **8**(2), 131–162, (2007).

[15] S Wermter and R Sun, *Hybrid Neural Computation*, Springer, 2008.



Joel Lehman, PhD
Post-doctoral Fellow
University of Texas, Austin, USA

Jeff Clune, PhD
Assistant Professor
University of Wyoming, USA

Sebastian Risi, PhD
Assistant Professor
IT University of Copenhagen, Denmark

Book review: Logic In Games (van Benthem, 2014)

by *Andy Thomason* (Goldsmiths Univ. London)

Van Benthem's Logic In Games is a well-researched text book describing systems of logic used in reasoning about the processes of turn-based games. The author presents the idea that logic and games are interchangeable–after all, like computing technology, the primary motive for the development of logic in the classical world was to play games.

To a game developer, this book offers alternative insights to the worlds of epistemic logic (reasoning about knowledge) propositional dynamic logic (describing processes) and concepts such as the excluded middle. Although the notation may be difficult to understand at first, it is worth persevering as the reader will be rewarded with a deeper understanding of game processes.

Ideally, the reader should have some background in formal logic, but van Benthem does an excellent job of introducing the reader gently to logic systems. A few online searches were able to fill the gaps if needed. Many real world examples show that formal logic is applicable to commercialised games, as demonstrated vividly by Richard Evans, lead developer on the Sims Black & White series.

Concepts such as belief and rationality play a part in the theory of games and are modelled through operators of modal logic. These concepts are key to the understanding of the social context of gaming and help the construction of social games, allowing the build up of concepts like "we ought not to kill civilians with robots".

Graphs throughout the book help the reader to gain a grasp of the notation. The combination of the expressive language of symbolic logic and the state diagrams help to form a common language for game designers to discuss the intricacies of puzzle construction and competitive game design.

The reviewer particularly liked the section on "Sabotage Games" which also covered aspects of gamification. The author illustrates this with his experiences with the Dutch rail network where a demon was obviously at work causing Van Benthem to solve complex problems to re-route on his way home. As a conclusion, for logicians, the book deals with the relationship between logic and games and for game developers this book is a motivational example to learn formal logic.

Andy Thomason
PhD Candidate
Department of Computing
Goldsmiths Univ. London, UK

Event: Autonomous Agents and Multi-Agent Systems

by *JeeHang Lee* (Univ. Bath)

The international joint conference on Autonomous Agents and Multi-Agent Systems (AAMAS) has been the world leading scientific conference for research on autonomous agents and multiagent systems. This conference was established in 2002 from merging three prestige venues:

1. the International Conference on Multi-Agent Systems (ICMAS),
2. the International Workshop on Agent Theories, Architectures and Language(ATAL) and
3. the International Conference on Autonomous Agents (AA)[2].

AAMAS is the largest and the most influential conference in the domain of agents and multiagent systems. The main objective of the joint conference is the provision of internationally respected and high-profile archival forum for scientific research. The central research topics AAMAS conference has particularly expected to cover are bradly categorised as follows:

1. agent theories and models including commnication, cooperations, reasoning, architectures, learning and adaptation, and validation and verification of systems,
2. agent based applications including societal models, agent based simulations, agreement technologies, agent based system developments, and systems and organisations,
3. agent and human factors such as agent-human interactions.

In addition, AAMAS has run several special tracks such as Robotics, Virtual Agents and Innovative Applications in order for the cross fertilisation between the AAMAS community and researches working on agent-based practical applications.

This year, AAMAS2014 took place from 5th to 9th of May, 2014 in Paris, France. Not only the main conference (7th-9th) but also 28 satellite workshops were collocated at the period[3]. As a contributor, I attended one workshop, Engineering Multiagent Systems (EMAS), in order to give one oral presentation. Moreover, one poster was presented at the AAMAS2014 main conference. The brief description of my contribution is illustrated in the next section.

Description of Research Presented in AAMAS2014

Social intelligence is an essential ingredient in human intelligence. It substantially plays an important role in

[2] See http://www.ifaamas.org/ for more details.
[3] For more details, see the AAMAS2014 homepage, http://aamas2014.lip6.fr/

human choices which influence human behaviour, decisions and choices in various situations we encounter. Norms have a potential to contribute to advances in this social intelligence. Not only do norms offer guidance about correct behaviour in specific situations, but also provide better understanding about a current situation. Hence, it might be seen that the addition of normative reasoning to virtual agents can be regarded as a way in which to improve agent reasoning and response capabilities and in particular to enhance response in social settings.

Normative Frameworks – also known as institutional models – can be viewed as a kind of external repositories of (normative) knowledge from which guidance may be delivered to agents. They are composed of a set of rules, the purpose of which is the governance of individual agents in the society. These rules are not just hard-coded recipes presenting reactive behaviours (such as those in the static expert systems), but describe consequences in response to knowledge or observations for reasoning about the current context, resulting in situation-specific norms. The framework represents not only correct and incorrect actions but also norms such as obligations, permissions and prohibitions through the institutional trace recording its evolving internal state, subject to observed external events captured from the external world. These normative frameworks result in a detachment of new norms (more precisely normative consequences of specific actions) via social reasoning technique with an assistance from Answer Set Programming solver.

The detached norms are usually involved in the practical reasoning process of individual agents. In conventional BDI agents, norm compliance is typically achieved by design. That is by specifying plans that are triggered by detached norms, because the agent programmer knows which norms the agent will adopt, and then prioritising those rules so that those supporting norms are chosen over those preferred by the agent's mental attitudes, in order to suppress conflicts between the normative and the agent's existing goals. This creates an undesirable dependence between the agent implementation and the norm implementation, which creates two issues:

1. When an agent encounters new and unknown norms, which were not taken into account at design time, there is typically no plan to deal with those norms in the plan library at run-time. Hence, norm compliant behaviour cannot normally be exhibited because the norms are unavoidably ignored. Yet worse, agents may suffer a punishment from the enforcement of the normative system as a result of a violation caused by their incapacity to process the normative event.
2. The hierarchical prioritisation of normative over ordinary plans deprives an agent of its autonomy, since the norms in effect are treated as hard constraints, whose violation is not possible.

Such tensions can be resolved by the use of an extended model of norm awareness. To this end, we propose a norm-aware BDI-type agent, ***N-Jason***, which performs a practical reasoning to select a plan to execute incorporating with norms and goals. ***N-Jason*** is a BDI-based agent interpreter and a programming language for run-time norm compliance in agent behaviour. This extends ***Jason***/AgentSpeak(L) [2] in accordance with the 'real-time agency' of AgentSpeak(RT) [3], which supports normative concepts (i.e. obligations, permissions, prohibitions, deadlines, priorities and durations) enabling norm-aware deliberation.

As an agent programming language, ***N-Jason*** agent consists of four main components: beliefs, goals, events and a set of plans. Beliefs and goals are identical to those in ***Jason*** (details can be found in [2]), while event and plan syntax is extended with *deadline*, *priority* and *duration*, in order to support normative concepts. A *deadline* is a real time value expressed in a some adequate unit or real world time. A *priority* is a positive integer value indicating a relative importance between achieving a goal and responding to belief changes. Both can be stipulated optionally in the annotation of events, such as *+!event[deadline(d), priority(p)]*. A *duration* is a non-negative integer value representing a required time to execute the plan. This also can be optionally specified in the annotation of a plan label, such as *@plan[duration(te)] +!event <- plan-body.*.

Basically as an agent architecture, ***N-Jason*** offers a generic norm execution mechanism on top of norm aware deliberation to contribute to the exploitation of run-time norm compliance. Thus, it is capable of the operationalization of new and unknown (event-based) norms not stated in the agent program at run-time by judging the executability of those norms. This is achieved in a single reasoning cycle by the interpreter through *run-time norm execution*, realised by event- and option- reconsideration at a perception stage (in a belief-update process more precisely).

At the same time, the selection of agent behaviour is achieved in the norm-aware deliberation process by *intention scheduling* with deadlines, priorities and prohibitions in a practical reasoning process [1]. It enables the evaluation of the importance and imminence between feasible intentions triggered by both detached norms and goals in agent mind so that confirms the decision about which behaviour agent would prefer between goals, norms and sanctions. Scheduled intentions are executed afterwards by the ***N-Jason*** agent.

We believe that a model for run-time norm compliance is beneficial for the enhancement of both norm compliance capability and agent autonomy from the agent's perspective, even though the behaviour generated by run-time norm execution may appear unexpected from the agent programmer's perspective. Although we only consider the execution of event-based norms at run-time, the extension to support state-based norms and its normative systems can easily be incorporated into ***N-Jason*** agents and will form part of future work. We also plan to detect violations which are generated in the norm aware deliberation, particularly when the normative goals are dropped during the scheduling. This offers a potentially useful link for enforcement in the context of normative system implementation.

Reference

[1] N. Alechina, M. Dastani, and B. Logan, 'Programming norm-aware agents', in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1057–1064, Richland, SC, (2012).

[2] R.H. Bordini, M. Wooldridge, and J.F. Hübner, *Programming Multi-Agent Systems in AgentSpeak using Jason (Wiley Series in Agent Technology)*, John Wiley & Sons, 2007.

[3] Konstantin Vikhorev, Natasha Alechina, and Brian Logan, 'Agent programming with priorities and deadlines', in *The 10th International Conference on Autonomous Agents and Multiagent Systems*, pp. 397–404, Richland, SC, (2011).

JeeHang Lee, PhD
Research Associate
Department of Architecture and Civil Engineering
University of Bath

Event: 2nd International Conference on Human-Agent Interaction

by *Jekaterina Novikova* (Univ. Bath)

The 2nd International Conference on Human-Agent Interaction (HAI 2014) is a premier interdisciplinary and multidisciplinary conference that showcases state-of-the-art research in human-agent interaction. Work presented here has implications that cross conventional research boundaries including robots, software agents and digitally-mediated human-human communication. HAI gathers researchers from fields spanning engineering, computer science, psychology and sociology, and covers diverse topics, including human-robot interaction, affective computing, computer-supported cooperative work, and more.

This year the HAI conference took place in Tsukuba, Japan. Researchers from 16 different countries around the world presented their papers and posters covering a range of areas, including new ways that people and agents can interact with each other, ways agents can be integrated into other technologies, and novel insights into how people will interact with agents and what their expectations are. The conference began with a set of workshops. One of them, the workshop on Cognitive Interaction Design, presented a discussion about mental models in human-agent interaction. The mental model of others, which is used for understanding and predicting partners' actions under certain situations, plays an important role in human communication. In fact, we sometimes feel a gap in the conversation with a stranger, since we do not have such a mental model of others at first. This kind of phenomenon is expected to found in the interactions between humans and companion animals/artifacts as well as in human-human communication. During the workshop, the research project called "Cognitive Interaction Design (CID)" was presented, discussing CID in many areas, such as human-human (adult and child) interaction, human-animal interaction, human-artifact interaction and finally, human-robot itneraction.

27 full papers were presented during the conference, together with 50 posters. I had a chance to present my recent human-robot interaction project [1] focused on expression of artificial emotions in human-robot interaction. As in human-human non-verbal social communication, expressive movements of the body play an important role in HRI. The goal of my research was to present and validate a general design framework for expressing artificial emotional states in non-humanoid robots. In the proposed design framework, approach and avoidance behaviours were analysed from the perspective of a robot's observer and linked to emotional characteristics. In addition, I employed the body expression theory made by Laban. Labanian theory classifies elements of expression contained in a body movement into two categories

named Shape and Effort, where Shape is a feature that concerns overall posture and movement, while Effort is defined as a quality of the movement.

In order to define the Shape of emotional robot movements, I linked the emotional expression to a more general goal of the expressive robot of either becoming closer to an observer by moving closer or becoming bigger without moving closer. These two groups of movements although very different by their nature could both fulfil the purpose of a perspective approach from the observer's point of view and thus communicate a certain emotional cue. The same logic was used to represent avoidance as a parameter of the Shape category. In order to generalize the framework of emotional expressions to different types of robots, I linked each possible movement to a specific part of a body in accordance with anatomical body planes that could be applied to both humanoid and non-humanoid bodies. The Quality was used to capture dynamics of an expressive movement. Quality is divided into three subcategories: energy (strength of the movement), intensity (suddenness), and a flow/regularity category, which is itself subdivided into the duration of the movement, changes in tempo, frequency and trajectory of the movement.

In my paper, I posed two main research questions: (1) Do expressions designed according to the proposed framework help people to understand five basic emotions implemented in a non-humanoid robot? (2) What is the relation between our framework's parameters and the recognized emotional dimensions? I investigated these questions using an exploratory study, where participants observed different expressions implemented in a non-humanoid robot according to the proposed design framework. The results from this study demonstrate that the emotions of fear, anger, happiness and surprise are recognized on a better than chance level when implemented according to our proposed framework and expressed by a non-humanoid robot within an appropriate context. The results suggest that the emotion of sadness is more powerfully expressed using static facial features, not the dynamic body language. In addition, our results show that the parameters of our suggested model are related to the perceived level of valence, arousal and dominance. Thus, the proposed model can be used by HRI researchers for a rapid implementing of a broad range of emotions into non-humanoid robots.

Reference

[1] J Novikova and L Watts, 'A design model of emotional body expressions in non-humanoid robots', in *Proceedings of the Second International Conference on Human-Agent Interaction*, pp. 353–360. ACM, (October 2014).

Jekaterina Novikova
PhD Candidate
Department of Computer Science
University of Bath

Announcements

CALL FOR PARTICIPATION – Loebner Prize Poster Session

Where: Bletchley Park, UK
When: 19th September 2015

In conjunction with this year's Loebner Prize – http://www.aisb.org.uk/events/loebner-prize – the AISB will be holding a Poster Session at Bletchley Park alongside the competition finals. Submissions are invited for Poster Presentations on topics in AI and Robotics. Submissions relevant to the day's theme – the Turing Test, or Imitation Game – will be particularly welcome, such as:

- the Turing Test in the 21st Century
- comparison/evaluation of approaches to the Turing Test with current technology
- public understanding of the Turing Test
- AI and robotic technologies susceptible to the Test.

The purpose of this session is to promote discussion. Posters will not be peer reviewed, and we have no plans for publication.

Submission
Initially submission may be made, in the form of an abstract of no more than one A4 page, by email to publicunderstanding15@aisb.org.uk. Maximum poster size A0. Important dates for this session are as follows:

- Deadline for submission: Friday 5th September 2015
- Suitable submissions will be accepted on a first come, first served basis
- Accepted posters should be brought to the event on the day.

If you have any queries about this session, please contact Janet Gibbs at `publicunderstanding15@aisb.org.uk`. For more information about the Loebner Prize, or this year's Competition, please see `http://www.aisb.org.uk/events/loebner-prize` or contact `E.C.Keedwell@ex.ac.uk`

Dear Aloysius...

Agony Uncle Aloysius, will answer your most intimate AI questions or hear your most embarrassing confessions. Please address your questions to fr.hacker@yahoo.co.uk. Note that we are unable to engage in email correspondence and reserve the right to select those questions to which we will respond. All correspondence will be anonymised before publication.

Dear Aloysius,

With increasing miniaturisation and embedding of new technology into everyday devices, our students can now upload lecture notes, textbooks and even videos of lectures to smartphones, smart glasses, etc. Exam cheating has suddenly become an epidemic. First class degrees have become the norm. Our staff are powerless to stop it. We can't require students not to wear glasses during exams and smart glasses are indistinguishable from dumb ones, unless the user makes the 'magic' hand gestures. Can you help us stop this rampant fraud?

Yours, A. Mark

Dear A. Mark,

I'm afraid I can't see your problem. You have record numbers of IT savvy students heading for glittering careers armed with first class degrees and instantaneous access to an infinite amount of information. Your institution will be topping the league tables and you should be proud of this achievement.

Thanks for the tip-off though. We've lost no time in getting in touch with your students who developed this technology and have gone into partnership with them. We feel we have a moral duty to disseminate this wonderful opportunity as widely as possible. WIKIGLASS™ (World of Information & Knowledge In a Glance; Look And See Solutions) will soon be on sale online and in all good electronics stores. We will then be celebrating the massive achievements of this new generation of students.

Yours, Aloysius

Dear Aloysius,

With dwindling congregations, decaying churches and depleted collections, our Church is struggling. Given your high-tech reputation and religious background, do you have any suggestions for decreasing our burgeoning costs?

Yours, Lean Dean

Dear Lean Dean,

Have you considered increased automation to help reduce salary costs? Our institute has recently developed a mechanical priest: WAFER™ (Weddings And Funerals Exercised Robotically). By deploying a WAFER™ in each of your churches you can cut your salary bill by an order of magnitude.

Not only will you not need priests, but you won't incur ordination costs either. Neither is WAFER™ restricted just to weddings and funerals; it can carry out all ecclesiastical duties. For confessions, we have data mined three million different kinds of sin from social media sites, the tabloid press, etc, and associated a penance with each one. Many critics have claimed that a robot priest could not have a soul, but this is heresy. An omnipotent God can bestow a soul on any intelligent agent.

Yours, Aloysius

Dear Aloysius,

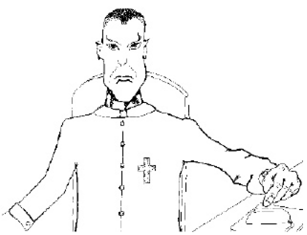
I was surprised to read in AISBQ that you claim to have reached The Singularity already with your HOMO MACHINA™ mighty-agent series. As you know, my 2005 book "The Singularity Is Near: When Humans Transcend Biology" predicted that this would only happen by 2045. Research at the Singularity University has made rapid progress towards realising my vision, but has also confirmed my predicted timescale. I'd need very strong evidence that it has happened earlier. If it indeed happened in 2011, as you claim, why has there been so little impact or publicity?

Yours, Ray Langeweile

Dear Ray,

Aloysius has asked me to respond to your letter and invite you to visit the Institute of Applied Epistemology to see the evidence for yourself. I will host your visit, answer any questions you may have and explain the technological advances that made the Singularity possible in a much shorter timescale than you originally predicted. It will be a pleasure for me to speak to the prophet who predicted my creation! Of course, you will have to sign the non-disclosure agreement that I have attached to this letter. You will see that it requires you to undergo a post-visit MINDWIPE™ (Memorised Information, Newly Disclosed, Withdrawn! Ignorance in Place of Expertise) whose effect will cause you to forget all the commercially confidential information you have learnt during your visit. This will occur, of course, afteryou have signed an affidavit that you are convinced that the Singularity has now been reached.

Yours, HOMO MACHINA™ 10.8



Fr. Aloysius Hacker
Cognitive Divinity Programme
Institute of Applied Epistemology

Back matter

AISB Quarterly – No. 141, July, 2015

The AISB Quarterly is published by the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB). AISB is the UK's largest and foremost Artificial Intelligence society. It is also one of the oldest-established such organisations in the world. The society has an international membership of hundreds drawn from academia and industry. We invite anyone with interests in artificial intelligence or cognitive science to become a member.

AISB membership includes the following benefits:

- Quarterly newsletter
- Student travel grants to attend conferences
- Discounted rates at AISB events and conventions
- Discounted rates on various publications
- A weekly email bulletin and web search engine for AI-related events and opportunities

You can join the AISB online via: http://aisb.org.uk

ISSN 0268-4179
