AISB Quarterly

No.127

The Newsletter of the Society for the Study of Artificial Intelligence and Simulation of Behaviour

Analyzing Traces of Activity for Modelling Cognitive Schemes of Operators

Modern design of human/machine interfaces requires a better understanding of how operators control their interaction with machines. To understand these interactions, cognitive ergonomists seek to construct cognitive models of operators. These models generally depict operator activity as a process of information-collecting, computing, decision-making, and action. While this symbolic approach effectively describes formal reasoning, it becomes ambiguous when considering an activity in which operators are physically involved, such as driving a car. Here, operators' cognitive process *accompanies* their actions and can be equally viewed as a cause or as a consequence of their activity. Perception, cognition, and action can hardly be separated, because expectations drive perception, and the feeling of comprehension relies on possibilities of action.

Where interaction and perception are so tightly coupled, we take inspiration from psychologists like Piaget, who have proposed to keep perception and action embedded into *schemes*. We consider schemes and *cognitive schemas* as the basic elements of our cognitive modelling, and we seek to highlight and model them from "traces of activity" (Georgeon, 2008). To do this, we have implemented *knowledge engineering* software and a method of cognitive modeling, which derives from "traces of activity". This software includes graph processing and visualization, symbolic inference, as well as ontology manipulation (Georgeon, Mille & Bellet, 2006).

The "traces of activity" are a sequence of events that describe the interaction of the driver with their environment. In our case, the trace gathers data describing the driver's behavior and situation: steering angle, pedal use, GPS positioning and cartography, distance ahead, and eye information. The trace also includes subjective evaluations made by the driver or by the researcher during the experiment, or during retrospective verbal protocols with video played.

The outline of the modeling process is given by figure 1, overleaf.

The activity over time is represented on the vertical axis. The curves symbolize the continuous flow of collected data. The horizontal axis represents the level of abstraction. The diagonal arrow represents the modeling process. Step 1 is data collection, while Step 2 consists of identifying the first level of points of interest. These points of interest are then processed by the system as symbols. Step 3 consists of inferring more abstract symbols from the basic symbols, and organizing them in an ontology. Step 4 consists of producing models of the activity on the basis of these symbols.

The *points of interest* and symbols are not found blindly by algorithms, but we specify them by looking at the data. They are points that interest us because they describe the activity in a way that helps us understand it better. Thus, we emphasize the interactivity of our software. These points are essentially defined on an evolutionist and pragmatic basis, i.e. trying to keep the most useful/meaningful point types. Once these types are specified, we make programs to identify their instances automatically in the trace.

The ontology supports the visualization parameters such as the symbols' color and shape. It also supports the semantics on which inference rules are based. Inference rules are a way to add new symbols in the trace. These new symbols represent more abstract concepts, which summarize patterns of lower level symbols. We thus construct a language for describing this activity.

Figure 2 shows an example of plot that we obtain, representing a motorway lane change (Henning, Georgeon & Krems, 2007). It shows a typical driving situation, where a slow vehicle impedes a driver. The driver may check his or her left mirror several times. Deciding to overtake the slower vehicle, the driver accelerates while simultaneously checking the mirror. If the left lane is clear, he or she switches on the blinker, starts steering, and crosses the line. The circles at the bottom represent low-level events.

In this issue:

Features

Reviews

Society

Father Hacker

The AISB Quarterly is published by the Society for the Study of Artificial Intelligence and Simulation of Behaviour.

ISSN 0268-4179

Editor
Colin G. Johnson
Computing Laboratory
University of Kent at Canterbury
aisbq@aisb.org.uk
http://www.aisb.org.uk/aisbq/index.shtml

Advertising and Administration
Therie Hendrey-Seabrook
aisbq@aisb.org.uk
School of Science and Technology
University of Sussex
Falmer, Brighton
UK, BN1 9QH
Tel: (01273) 678448
Fax: (01273) 671320

Analysing traces of activity for modelling cognitive schemes of operators (continued)

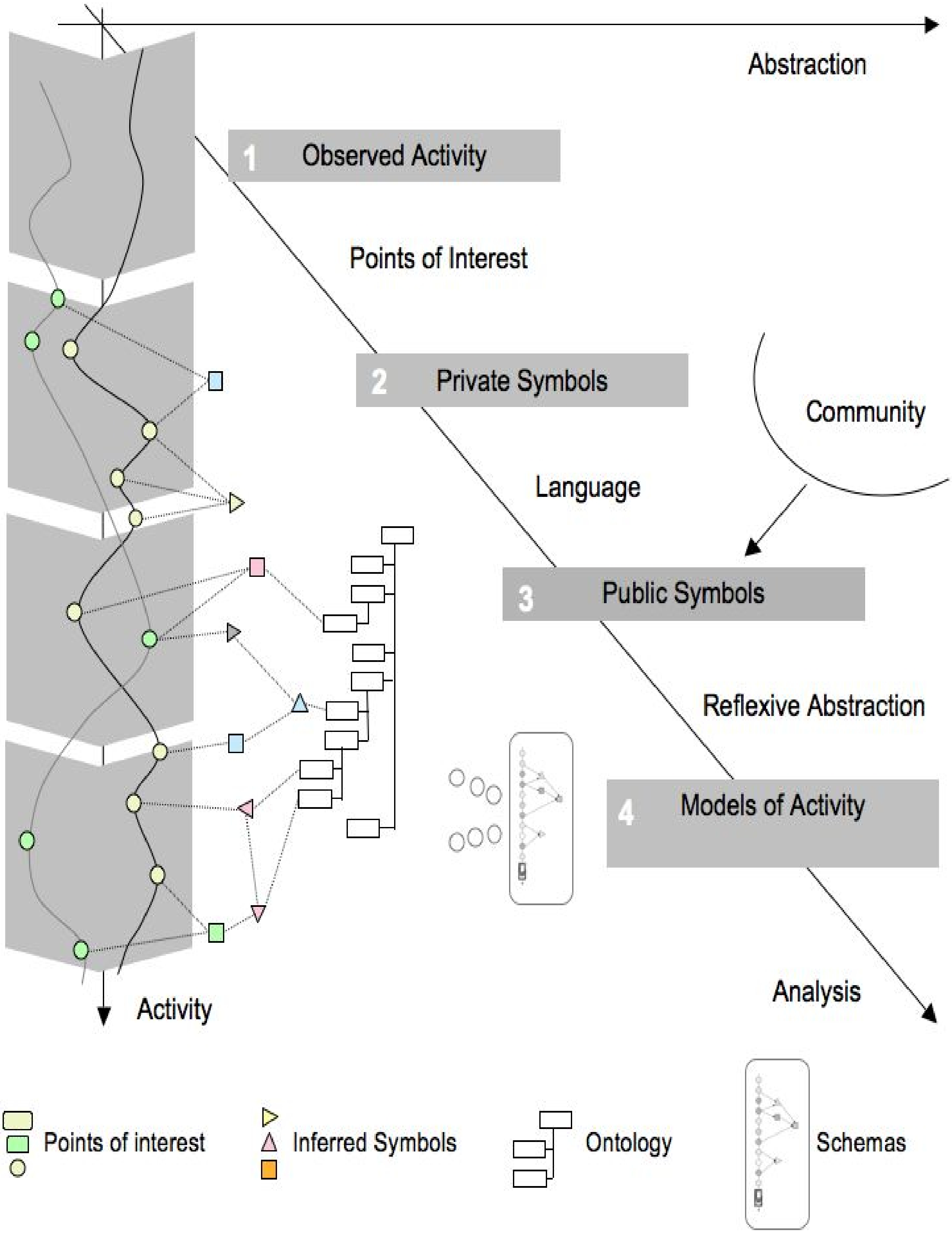


Figure 1: Process of Analysis

Analysing traces of activity for modelling cognitive schems of operators (continued)

The upper part represents the high level symbols. Lines between them represent inference relations from lower to higher. Longitudinal information is represented on the axis, things concerning left are above, and right are below. In this situational category, the conjunction of acceleration and left mirror glance indicates the decision to overtaking the impeding vehicle. From this, we can compute a "marker" of the decision (violet triangle at -3s). It occurs about one second before the blinker is switched on — it is thus a predictor of the maneuver. As ergonomists, we explain this pattern of behavior as the performing of a cognitive schema adapted to a category of situation, that we classify in parallel. It involves unconscious know-how, connected to some points of decision at a more conscious level.

From an epistemological point of view, our approach lets us connect a bottom-up with a top-down modeling process, i.e. connecting experimental data with psychological explanations. We offer pragmatic arguments in support of cognitive schemas as a means of explaining how humans perform their activities. Our approach is based on a *constructivist* epistemology, since models are built through an evolutionist and pragmatic process, and driven by mindful analysts. We claim that this process can provide insights about how salient events of activity can arise into consciousness and become the basis for symbolic reasoning. This leads us to propose it as a "constructivist model of awareness".

References

Georgeon O., Mille A. & Bellet, T. (2006). Analyzing behavioral data for refining cognitive models of operator. Philosophies and Methodologies for Knowledge Discovery, Seventeenth international Workshop on Database and Expert Systems Applications (pp. 588-592). Krakow, Poland: IEEE Computer Society Press.

Henning M., Georgeon O. & Krems J. (2007). The quality of behavioral and environmental indicators used to infer the intention to change lanes. International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. Stevenson, Washington. July 9-12, 2007.

Georgeon O. (2008). Analyse de traces d'activité pour la modélisation cognitive: Application à la conduite automobile. PhD Thesis. Université Lyon 2. Lyon, France.

Olivier Georgeon
Pennsylvania State University
Olivier.Georgeon@gmail.com

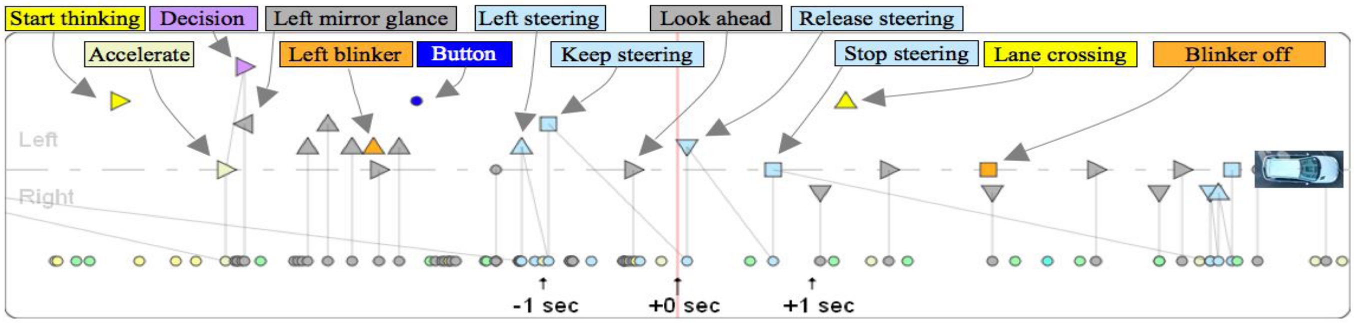


Figure 2: Motorway Lane Change with Acceleration.

Foundations of a Philosophy of Collective Intelligence

Philosophy, artificial intelligence, and cognitive science have long been dominated by the presupposition that intelligence is fundamentally individual. More recent work in tracking the behavior of individuals finds that their behavior - ranging from movement to turn-taking in conversation - can be reliably tracked by appealing to the behavior of others in their social network and their immediate environment with a high degree of accuracy (over 40 to 80% of variation over a wide variety of tasks) without any appeal to planning, reasoning, or verbal language (Pentland, 2007). Increasingly, intelligence is seen not as having its locus in the individual, but in the network of relationships that the individual has with the external world and other individuals.

Let us consider this philosophically from the perspective of cognitive science. In their Extended Mind Hypothesis, Clark and Chalmers introduce us to Otto, a man with an impaired memory who navigates about his life via the use of notes in his notebook (Clark and Chalmers, 1998). Otto wants to navigate to the Metropolitan Museum of Modern Art in New York City from his house in Brooklyn, but to do so with his impaired memory he needs at least the address; he needs a map. The map is just an external representation in the environment of Otto, and can drive the cognitive processes of Otto in a similar fashion to the way that classical AI assumed internal representations in Otto's head did. Clark and Chalmers point out if external factors are driving the process, then they deserve some of the credit. In this regard, the Extended Mind thesis undermines the strict division between internal and external of the agent itself.

Imagine the world to be inhabited by multiple individuals that can access the same representation. In almost all the original examples used in the Extended Mind argument, they deploy a single person sitting in front of a computer screen. A more intuitive example would be two people using the Internet to both share a single representation. One could imagine Otto trying to find his way to the Museum of Modern Art, and instead of a notebook having a personal digital assistant with access to a map on the Web. Likewise another person, say Inga, can have access to the exact same map via her personal digital assistant. Since both Otto and Inga are sharing the exact same representation and because they are both using it in the same manner, Inga and Otto can be said to share at least some of the same cognitive state, due to the fact that their individual cognitive states are causally dependent on accessing the same representation. This representation is the "same" precisely because the digital memory of the computer allows near-perfect copies.

Unlike the lone digital computer, what the Web specializes in is allowing almost everybody to access the same set of representations simultaneously with increasingly low latency. This is the true cognitive revolution of the Web. The value of external representations comes with their accessibility, for an external representation that is not accessible when its needed cannot be used to enable online intelligence. The Web provides precisely the cognitive scaffolding to enable distributed individuals to rapidly co-ordinate in near real-time through the modifications of representation. In light of this, conservatively the study of external representations should have an increasingly large role to play in cognitive science. More radically, is there any difference that makes a difference between internal and external representations? If not, do we still have individuals in any useful sense of the term, if intelligence is increasingly outside the skin?

To overcome the individual-as-body-in-skin presupposition that is so heavily built into Anglo-American philosophy, what we need is not German philosophy, but French philosophy. French theorists Deleuze and Guattari put forward a concept that can replace the notion of a body: the assemblage. In contrast with the individual body, Deleuze and Guattari "call an assemblage every constellation of singularities and traits deducted from the flow - selected, organized, stratified - in such a way as to converge artificially and naturally" (Deleuze and Guattari, 1987). Defining intelligence in terms of a fully autonomous agent bound by the skin is not even an accurate portrayal of human intelligence, but a certain conception of the individual human subject, "a certain conception that may have applied, at best, to that faction of humanity who had the wealth, power, and leisure to conceptualize themselves as autonomous beings exercising their will through individual agency and choice" (Hayles, 1999). By jettisoning this conception, and maintaining the commitment to a certain necessary degree of embodiment, cognitive science can do justice to complex phenomenon such as the advent of the Web and increasing recognition of the collective nature of intelligence. The vast technological changes humanity has engendered across the world are now reshaping the boundaries of human bodies, and so the cognitive world and the domain of cognitive science. This has been a process that has been ongoing since the dawn of humanity, but only now due to the incredible rate of technological progress, as exemplified by the growth of collective intelligence on the Web, does it become self-evident.

References

A. Clark and D. Chalmers, The Extended Mind, Analysis 58, p. 10-23, 1998.

G. Deleuze and F. Guattari, A Thousand Plateaus, University of Minnesota Press, Minneapolis,1987.

N. Hayles, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics, University of Chicago Press, Chicago, 1999.

A. Pentland, On the Collective Nature of Human Intelligence Adaptive Behavior 15(2), p. 189-198, 2007.

Harry Halpin
University of Edinburgh
H.Halpin@ed.ac.uk

An extended version of this article can be found the the AISB'08 Conference Proceedings.

Using Recency and Relevance to Assess Trust and Reputation

Interactions among software agents occur when individual agents are unable to achieve their goals alone. In decentralised multi-agent systems, agents possess limited information about their environment and other agents. Additionally, agents are assumed to be self-interested and cannot be expected to always be truthful about their performance and opinions. Each agent aims to its individual goals, and the choice of who to interact with determines, in part, whether goals are fully achieved, partially achieved, or fail. Therefore, the selection of appropriate and reliable interaction partners is crucial to the success of subsequent interactions.

In open and dynamic environments the selection of interaction partners is made more complex due to potential rapid and unforeseen changes in agent behaviour and the agent population itself. Our research aims to minimise the risk associated with the uncertainty of agent interactions. We adopt the commonly used concepts of trust and reputation in managing this uncertainty. Trust is defined as the level of risk associated with cooperating with another agent and it estimates how likely the agent is to fulfil its commitments [3]. Trust can be derived from direct interactions among agents and from reputation, which is built from information received from third parties. Agents can thus make more informed decisions about whether to interact with others, based on trust and reputation. Figure 1 illustrates the type of agent interactions we consider. The evaluator is the agent assessing the trustworthiness of the target agent as a service provider for Task k. The dashed line indicates that there are no previous interactions between the evaluator and target for that specific task. The evaluator can request recommendations from witnesses. In this case, the witnesses have directly interacted with the target, as depicted by the solid lines, and I+ and I- represent the number of positive and negative interactions between two agents.

We propose a model of trust and reputation that enables agents to adapt quickly in dynamic environments by judiciously choosing interaction partners. Our approach combines components from several existing models, and we build upon these to more efficiently determine trust from direct experiences and recommendations. We take a

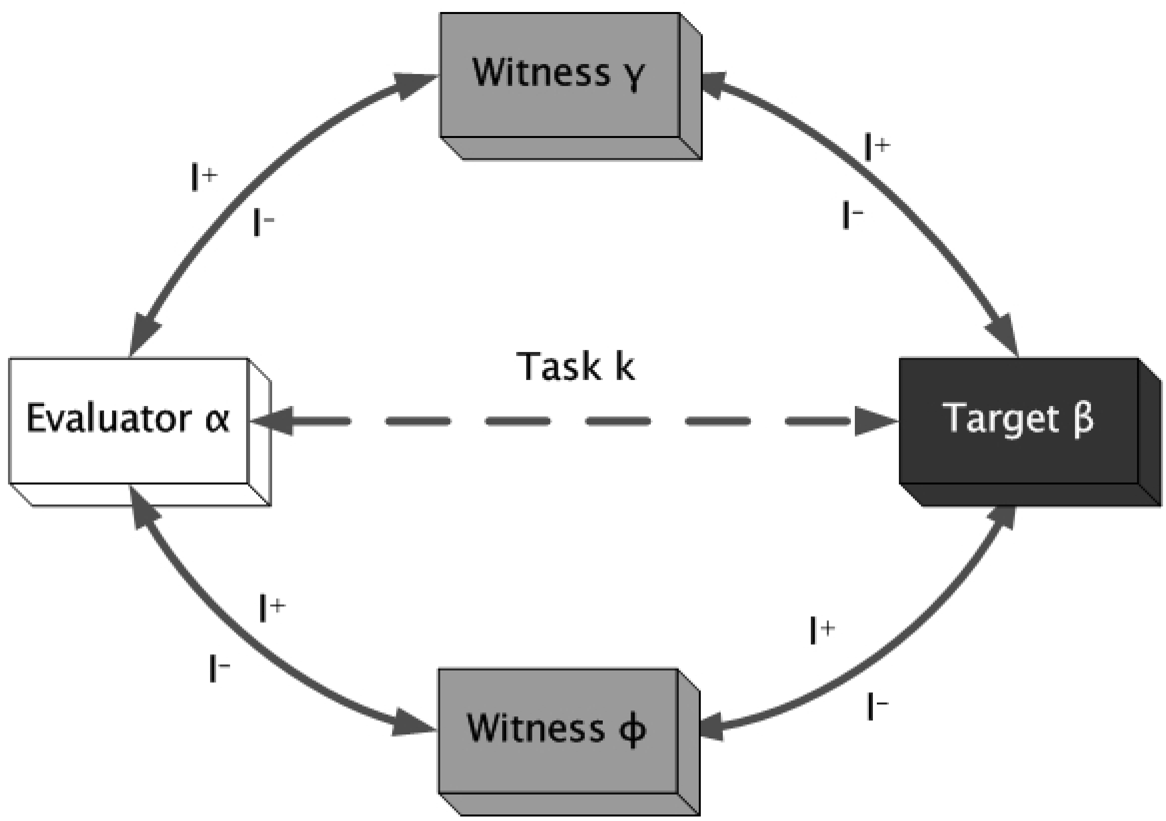


Figure 1: Illustration of agent interactions between the evaluator, target, and witnesses.

Using Recency and Relevance to Assess Trust and Reputation (continued)

multi-dimensional approach for evaluating trust and reputation and include a richer set of recommendation information sharing, to allow for the recency and nature of interactions to be communicated. Moreover, agents can better select recommenders based on the relevance of their opinions for the evaluating agent's purposes.

Some of the most relevant related work includes Marsh's trust formalism [7], which considers direct interactions among agents and divides trust into basic trust (representing an agent's trusting disposition), general trust (the perceived reliability irrespective of context), and situational trust (the reliability in a specific situation). We use these three views of trust and complement direct trust with witness reputation to achieve greater accuracy in predicting agent behaviour. ReGreT [8], FIRE [5], Ntropi [1,2], and MDT-R are approaches that also combine both trust and reputation. ReGreT uses three dimensions of information to assess reputation: individual (from direct interactions), social (from experiences of the evaluator's group with the target), and ontological (relating different aspects of reputation). FIRE integrates up to four information sources: interaction trust (from direct interactions), role-based trust (from social relationships), witness reputation (built from witness reports), and certified reputation (rating references provided via a trusted third party). Ntropi models trust and reputation with distinct levels and uses recommender trust to assess witness credibility. MDT-R considers multi-dimensional trust and recommendations where trustworthiness is modelled according to various criteria, such as cost, quality, and timeliness. This helps to maintain the original agent interaction data for the evaluator's own use and for information sharing.

Our model is broadly based on MDT-R and adopts the multi-dimensionality of trust and recommendations, as well as the sharing of interaction summaries. We extend MDT-R by including a richer set of information on recency and the experience of witnesses when sharing information. This allows the evaluator to more accurately select witnesses, and thereby providers. We use two main sources of trust information in our model, direct trust from interactions between the evaluator and the target, and witness reputation, which is built from direct and indirect recommendations from third parties. Trust is represented as a continuous value between -1 and +1, which allows us to maintain both sensitivity and accuracy. However, as in Ntropi, we use discrete levels to compare trust values to give a simple comparison method and avoid overfitting. Recommendation trust is used by the evaluator to assess the accuracy of witnesses in their recommendation. The evaluator uses this, together with witness experience and the relevance of opinions based on the recency of interactions to associate weight with recommendations. Reputation is then computed from the recommendations obtained, with more relevant recommendations having greater weight. Direct trust and witness reputation are aggregated to compute a performance value for each potential provider. Various factors contribute to this assessment, according to the importance of direct trust, witness reputation, and advertised service characteristics.

We have built a simulated environment to validate our approach. Initial experimental results demonstrate that the use of both trust and reputation in assessing trustworthiness can facilitate more effective interaction partner selection. We have compared the effectiveness of using trust, and trust with reputation, against using single service characteristics in a number of settings. From our experiments, we observe that using trust and reputation gives better results in most cases, compared to using service characteristics only. For instance, in an experiment where the population contains dishonest agents, the success rate was higher when trust or trust with reputation is used by an average of 18%, compared to using cost as service characteristic. However, further experimentation is needed to determine the circumstances in which the improvement of using trust with reputation is significantly better than using trust only. We have also considered how our model performs when agents change behaviour as our aim is to enable the evaluator to quickly adapt its strategy. We observe that evaluators reassess trust quicker when they store smaller interaction histories, but this can potentially be exploited by malicious agents. Future work will consider how agents can achieve a compromise between quickly adaptation and guarding against malicious behaviour. More details on our model can be found in our paper published in the AISB 2008 Symposium [6].

References

1. A. Abdul-Rahman, A Framework for Decentralised Trust Reasoning, PhD thesis, Department of Computer Science, University College London, 2005.

2. A. Abdul-Rahman and S. Hailes, Supporting Trust in Virtual Communities, Proceedings of the 33rd Hawaii International Conference on System Sciences, p. 6007, IEEE Computer Society, 2000.

3. D. Gambetta, Trust: Making and Breaking of Cooperative Relations, Oxford:Blackwell, 2000.

4. N. Griffiths, Enhancing Peer-to-Peer Collaboration Using Trust, International Journal of Expert Systems with Applications 31 (4), p 849-858, 2006.

5. T. D. Huynh, N. R. Jennings and N. Shadbolt, An Integrated Trust and Reputation Model for Open Multi-agent Systems, Journal of Autonomous Agents and Multi-Agent Systems 13 (2), pp. 119-154, 2006.

6. S. N. Lim Choi Keung and N. Griffiths, Using Recency and Relevance to Assess Trust and Reputation, Proceedings of AISB 2008 Symposium on Behaviour Regulation in Multi-Agent Systems, 4, pp. 13-18, SSAISB, 2008.

7. S. Marsh, Formalising Trust as a Computational Concept, PhD thesis, Department of Computer Science, University of Stirling, 1994.

8. J. Sabater and C. Sierra, REGRET: Reputation in Gregarious Societies, Proceedings of the Fifth International Conference on Autonomous Agents, pp. 194-195, ACM Press, 2002.

Sarah N. Lim Choi Keung
and Nathan Griffiths
Department of Computer Science
University of Warwick
slck@dcs.warwick.ac.uk

A Short Review of Theorem Proving

Whenever Sherlock Holmes solves a case in a Conan Doyle book, the suspect always seems to get killed trying to escape or else confesses to everything, handily supplying all the gory details. Just once I would like to see the suspect say "Alright, it's time to get the lawyers involved - I'll see you in court". If this happened, it's possible that the suspect would win, as the case put forward by Holmes would be quite unsound in places.

Holmes uses various forms of reasoning to solve his cases. These include *induction, abduction,* and occasionally his trademark: *deduction*. Induction is the process by which we generalise rules from examples. So, for instance, if a silk scarf was left at each of ten murder sites, Holmes might induce the fact that this is something to do with the murders themselves, and expect to see one at the site of the next murder. Abduction is the process whereby we assign explanations to observations. In the case of the silk scarf, based on previous experience, Holmes might abduce the explanation that the killer leaves the scarf as a calling card. Deduction is the process whereby we derive new facts from old ones in such a way that the new facts must follow from the old ones. In the case of the silk scarf, if Holmes narrowed it down to two suspects, one of whom had an iron-clad alibi, then he could deduce that the other suspect was the killer.

Induction and abduction are unsound procedures because they are based on guesswork. For example, if all the ten murder victims were members of the silk scarf appreciation society, then the scarves would possibly have no connection with the murderer. Deduction, on the other hand, is a sound procedure. If we know that some old facts are indeed true, then any new facts deduced from them are guaranteed to be true also. Deduction can be used opportunistically to find new facts from old ones. Often, however, we want to do things the other way around: we start with some things we know to be true (called axioms) and something we suspect to be true (called a conjecture). If we can show that the conjecture follows deductively from the axioms, then we have proved it to be true and we would upgrade the conjecture to a theorem. More importantly, because we have used deduction, we know that our reasoning is sound and that the theorem really is true. Getting computers to perform such theorem proving is a well established and very successful area of Artificial Intelligence.

Proving that something is genuinely correct is a very powerful tool to have in your toolbox. For instance, suppose that, through extensive testing, a company is absolutely positive that certain electronic components perform in specific ways. This information can be used as axioms. When they combine these components in a circuit board, however, things get more complicated. While the company might be quite sure that the circuit board works as they want it to, they would probably want proof of this. Such a proof can be deduced from the axioms about the individual components using automated theorem proving software. The verification of hardware and software is an important application of AI techniques.

In many cases, it wouldn't be too bad if the hardware or software didn't quite perform as specified. However, there are many situations which are safety critical, and it is absolutely essential to prove that the hardware and software perform as we want them to. Next time you are in an aeroplane, you can rest assured that the computers helping fly the plane are using hardware and software which have been automatically verified to perform as they should. The same is true for power stations, medical equipment, and even your humble home computer.

Automated theorem proving is enabled by our understanding of logic. The word 'logic' can be interpreted in many ways. When we talk about a particular logic, we are really describing a language which enables us to express certain things in a very constrained way. If you think of how many times you have misunderstood someone over something quite simple, you can see the advantage of having a very restricted language which everyone agrees upon. This is especially true of computers, as they aren't so good at interpreting sentences. There are many different logics that we can pick and choose from to fit the nature of the theorem we want to prove. The simplest ones such as propositional and first order logic have been studied for centuries and are extremely useful. However, they have the drawback that certain things we might want to say in English cannot be expressed in these logics. Other logics are said to be more expressive if we can say more with them. For instance, if we need to express the fact that certain things change over time, we would need a temporal logic. Similarly, if we need to express the fact that certain statements are probably true, we would need a probabilistic logic. Probably. If we restrict our language to a logic, we can employ rules to perform deduction. These are called rules of inference, because we use them to infer something new. The simplest such rule is called Modus Ponens. As an example of Modus Ponens in action, suppose Sherlock knows for sure that silk scarves can only be bought from one shop in London. On discovering the first scarf at a murder site, he can deduce that it must have been purchased from that shop - a valuable lead for him to follow.

There are many similar rules of inference and the goal in automated theorem proving is to find a route from the axioms to the conjecture using only such rules. There are many ways to do this, and there is continued research into making the process better. One important method is proof by contradiction: pretend that the conjecture is false and deduce something silly as a result (i.e., something which contradicts the axioms, which we know are true). If the axioms are contradicted, then the conjecture can't be false, so it must be true. Automated theorem proving has great potential for an area of human endeavour where deduction rules supreme: mathematics. Unfortunately, as with many things in AI, we underestimated how difficult it is to prove mathematical theorems, and AI has not had a big impact here. Very occasionally, though, theorem provers have beaten mathematicians. In particular, as reported on the front page of the New York Times, a theorem prover developed by researchers in Chicago managed to deductively prove the Robbins conjecture, which had eluded mathematicians for 70 years. Sherlock would be proud.

Simon Colton, Imperial College, London
sgc@doc.ic.ac.uk
http://www.doc.ic.ac.uk/~sgc/

Conference Report: 9th European Conference on Artificial Life

Started in 1991, the European Conference on Artificial Life (ECAL2007) had its 9th meeting held on September 10th to 14th 2007, in Lisbon, Portugal. The scope of Artificial Life (ALife), quoted from the ECAL2007 website, "ranges from the investigation of how life or life-like properties develop from inorganic components to how cognitive processes emerge in natural or artificial systems", "through methodologies of synthesis implemented in computational, robotic or other artificial architectures." Being a leading conference in the field, ECAL2007 has gathered together researchers worldwide, from various areas such as computer science, engineering, biology, physics, chemistry, etc. Besides workshops and tutorials, oral presentations and poster pitches, a panel discussion was organised every day after the conference finished. Other events took place throughout the city, such as the Music-AL concert, aiming to communicate ALife ideas to a broader audience. Overall, participants were able to present the latest advances and communicate to get inspiration. This report will firstly introduce you to a workshop that the author attended and then a few keynote lectures and presentations, through which hopefully you may get an idea of the conference.

The Music-AL workshop in ECAL2007 brought together musicians and researchers working with the artificial life approach to create and study music. The ALife approach to music provides composers with innovative ideas and means for generating music. And it also enables researchers to study or model musicology-related issues. Work presented at the workshop covered a broad scope including music tuning systems, composition and improvisation, music performance, sound synthesis, software tools, and so on. Some had applied common ALife algorithms or mechanisms such as multi-agent simulation, genetic algorithms, cellular automata etc to specific music problems. Others demonstrated fresh ideas concerning generative structure, mapping or processing in music production. By highlighting the developments, the attendees were assured that research combining ALife and Music is at an exciting stage. During the panel discussion, the participatants had a chance to share their views about the importance and difficulty of balancing artistic and scientific values in this area. Above all, the Music-AL workshop was encouraging and thought-provoking for all attendees.

During ECAL2007, the presentations and posters appeared to be categorized under a few keynote lectures, presented by well-known researchers from related fields. Rudolf Bannasch gave a very interesting talk about "morphological intelligence in bionic applications". He has shown several examples of construction inspired by the functional anatomy of living organisms, such as the self-adaptive behaviour of fish fins (Fin Ray Effect ®), the humanoid robot, and so forth. Peter Todd introduced the advantage of using artificial-life-inspired models in studying psychology and cognitive science. By showing the results from some agent-based modelling of search behaviour in mating, eating, and parking, he demonstrated that agents generate certain effective behaviours via taking account of environment, even through very simple mechanisms. Another informative lecture was given by Janet Wiles on using computational modelling to study complexity in biological systems. She described a multi-level approach to modelling different levels of biological complexity: nucleotide sequences, genetic networks and ontogeny. She not only discussed the biological perspective on what roles these layers play, but also summarized the software challenges inherent in the project. It was also a great opportunity to listen to immunologist António Coutinho introduce the development in studying on adaptive immune system, a helpful talk for researchers working on artificial immune systems and their applications.

ECAL2007 included much more than this short report can reflect. For those interested, please refer to its homepage http://www.ecal2007.org/. The next ECAL will take place in Budapest, 2009.

Qijun Zhang
Interdisciplinary Centre of Computer Music Research
University of Plymouth
Email: qijun.zhang@plymouth.ac.uk

Conference Report: Third International Conference on Cognitive Science

"Cognitive Technologies" is a relatively new term, the use of which is becoming quite widespread in the cognitive science community as well and has a multidisciplinary nature. The term also refers very closely to an emerging interdisciplinary discipline of cognitive informatics (CI) that studies the natural intelligence and internal information processing mechanisms of the brain, as well as the processes involved in perception and cognition. According to the wiki definition, "CI provides a coherent set of fundamental theories, and contemporary mathematics, which form the foundation for most information and knowledge-based science and engineering disciplines such as computer science, cognitive science, neuropsychology, systems science, cybernetics, software engineering, and knowledge engineering" (URL 01). If we google the phrase "cognitive technology", there are now 386,000 occurrences of it. It struck me particularly while I was attending the presidential address of Boris Velichkovsky (2008), the chairperson of the Third International Conference on Cognitive Science, held form June 20-25, 2008 in Moscow, Russia. Two other papers presented at that conference that I would particularly like to mention in the context of this area are those of Dascal (2008) and Wachsmuth (2008). I particularly chose these two papers here because there is a vast contrast in the context in which the term "cognitive technologies" is used. The former is has more philosophical approach whereas the latter is on track of artificial intelligence and agent technology.

Marcello's paper on "Dialectical cognitive technologies," looks at language as a cognitive technology. The hypothesis suggested in the paper is that, "the very existence of this linguistic mode of communication and its set of tools acts as a cognitive technology in terms of which we tend to conceptualize controversial issues and take decisions, even in the absence of an actual adversary to face and defeat. (Dascal, 2008, p 35)" The author posits that it needs to be answered, with the advancements in the computer and information technology, whether or not, to what extent, and which of the language-based cognitive technologies we naturally use can be emulated by the kinds of technologies presently or in the foreseeable future available (Dascal, 2004).

Wachsmuth and Colleagues work is based on AI and agent technology. It surveys the recent research at Bielefeld University, in the department of AI, on "virtual humans with affective minds." It describes how an agent "Max", built using a cognitive architecture incorporates not only modules for perception, action, cognition but also emotion. Max has been deployed in various real life settings and games.

It is interesting that "Cognitive Technologies", the series, complements the series Lecture notes in Artificial Intelligence (LNAI) which encompasses artificial intelligence and its subfields and related areas, such as natural-language processing and technologies, high-level computer vision, cognitive robotics, automated reasoning, multi-agent systems, symbolic learning theories and practice, knowledge representation and the semantic web, and intelligent tutoring systems and AI and education. I would also like to mention a recently published book by Elsevier, titled, "Cognitive Technology, In Search of the Humane Interface," edited by B. Goraska and J.L Mey. This book is composed of a carefully gathered contributions of explorations of the human mind via the technologies the mind produces.

In the end, I feel I can tie this up with my area of research interest, which is related to cognitive modeling of attention (Hussain & Wood, 2008). I use hybrid cognitive architectures like ACT-R6 (Anderson et al., 2004) which has both symbolic and subsymbolic constructs and may also fall under the umbrella of "Cognitive Technologies."

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. Psychological Review 111 (4). 1036-1060.

Dascal, M. Dialectical Cognitive Technologies. (2008). In Proceedings of the third international conference on Cognitive Science, Moscow, Russia.

Dascal, M. Language as a cognitive technology, (2004). International Journal of Cognition and Technology 1(1): 35-89, 2002. (reprinted in B. Gorayska and J.L. Mey (eds.), Cognition and Technology: Co-existence, Convergence, and Evolution. Amsterdam: John Benjamins, pp. 37-62..

Hussain, F. & Wood, S. A cognitive model of Modulation between Attentional Networks, (2008). In Proceedings of the third international conference on Cognitive Science, Moscow, Russia.

URL 01: http://en.wikipedia.org/wiki/Cognitive_Informatics

Wachsmuth, I. Virtual humans with Affective Minds: Recent Research at Bielefeld (2008). In Proceedings of the third international conference on Cognitive Science, Moscow, Russia.

Velichkosky, B.M. Studying Cognitive Systems and Developing Cognitive technologies (2008). In Proceedings of the Third International Conference on Cognitive Science, USTC Press, Beijing Moscow, Russia.

Fehmida Hussain
University of Sussex
F.Hussain@sussex.ac.uk

Conference Report: ASSC-12, Taipei

The recent conference in Taipei of the Association for the Scientific Study of Consciousness (ASSC) was a fantastic opportunity to meet a large number of Asian postgraduates I would probably have never met otherwise – some of whom may subsequently be lured to next year's ASSC conference, in Berlin. It was also an opportunity to revisit a number of themes of personal interest. Although the neuroscience talks outnumbered the philosophy talks (roughly reflecting the difference in membership numbers within the ASSC), as a philosopher I will focus on the philosophy.

First, although it has not become central in my thesis (which is attempting to develop an enactive theory of concepts), I have a very strong intuition that our multiple concepts of self play a critical role in the coordination and structuring of our conceptual thought. The concepts of self is both like and in fundamental ways unlike the other concepts we possess; and, perhaps, the ways in which it is different from other concepts helps explain the role that it plays.

Thomas Metzinger's keynote address helped me understand much better what Metzinger means when he says things like no one has or has ever had a self. I'm still not sure to what extent I agree with him (e.g. whether "ontologically, no such things as selves exist"), but his central point is well taken: that our sense of self as a stable, even static, entity cannot possibly match the underlying reality: to the extent we depend on such a conception, there is the appearance of a necessary fiction, but a fiction nonetheless. If there is any basis to an ontological self – and I suspect there is – it's not at all what the phenomenal self would suggest.

So on the one hand, I'm not convinced that a first-person perspective is quite the "vague metaphor" that Metzinger takes it to be: difficult to describe in non-subjective, non-first-person terms, sure, but isn't that the point? I suspect that most of us really are quite clear what we mean by "I", even if we can't articulate it, even to ourselves. On the other, I think Metzinger is very close to being right in the things he says about the role the self plays in providing a sense of coherence and unity. I am encouraged as well that Metzinger takes phenomenal self as a serious matter to be explained and not simply explained away.

On a related note, Kristina Musholt gave a poster presentation – one poster board over from my own poster presentation – on the conceptualization of self and the relationship between self-consciousness, non-conceptual content and intersubjectivity. She asks: at what point does a self (implicitly required in any consideration of conscious experience) become aware of itself as a self? Among her conclusions: having a first-person perspective is not the same as being self-conscious (indeed, is prior to); and being self-conscious arises (can arise?) only in a social context: i.e., subjective self-consciousness depends upon intersubjective experience.

Certainly I'm inclined to agree that, at the very least, the social dimension transforms both our concept and our experience of self. Further I agree with what I take to be her implicit conclusion, that there are multiple concepts of self relating to different levels of abstraction away from the underlying fundamental nature of the organism: the self as the physical organism, the self as the experiencing (mental) entity, the self as myself, and so on, all of which have an unfortunate tendency to get conflated.

Second, concepts are often if not typically described as mental representations, and so I have the frustration, as Nicholas Georgalis gave expression to in his talk on Cognition and Consciousness, that the term "representation" is often used without being defined, and that what usually should be called representations get conflated with, to use Georgalis' terminology, "information-bearing states". If being a representation requires an agent using the representation as a representation – if it requires intentionality – then talk of unconscious or sub-personal representations risks confusion at best and incoherence at worst. I would go further, I think, than Georgalis, and, in the terminology of Sussex University's Inman Harvey, define representation as a four-part relation of *p* using *q* to represent *r* to *s*, where *p* and *s* may be the same or different agents. I also would be more willing than Georgalis to attribute representational states to certain non-humans; although the lack of linguistic medium prevents us from confirming these representational states directly, there are some ingeniously designed experiments (e.g., with the parrot Alex) that, I believe, make the attribution of representational states a fairly safe conclusion.

Third, there is the question of phenomenal concepts, which some philosophers, including ASSC speaker Yasuko Kitano, would use to address the so-called explanatory gap: the difficulty of accounting for phenomenal experience in purely objective, third-person terms. Phenomenal concepts are concepts of things in experience and so to be distinguished from concepts of, e.g., things in the world. Kitano considered, and rejected, such well-known objections to phenomenal concepts as Jesse Prinz's claim that the explanation offered by phenomenal concepts can better be explained in terms of non-conceptual "mental pointing". In the end I remain unconvinced whether phenomenal concepts can do the job that is required of them, though hopefully I have a better grasp of why some find them so intuitively appealing.

Joel Parthemore
University of Sussex
J.E.Parthemore@sussex.ac.uk

Society News

AISB Symposium on the Turing Test

- Sunday 12th October, 2008
- Palmer Building, University of Reading, Whiteknights, UK.
- Sponsored by the AISB, University of Reading and KYBERNETES

To coincide with the annual Loebner Prize this year to be held at the University of Reading (UK) on the 12th October, the AISB has also elected to sponsor a small one day invited-speaker symposium to present an alternative, formal, academic critique of issues surrounding the Turing Test (TT), one of the aims of which is to attempt to clarify two core issues surrounding the TT: (a) is a canonical reading of Turing's Test possible and if so (b) what might such a canonical reading be?

The day will commence with a selection of speakers offering a context and outlining a special perspective on the TT. These presentations are to be backed by four more focussed talks addressing specific issues related to the Turing Test (e.g. definitional; adequacy; tests in other modalities; technical/computational issues). As added relevance, the event is scheduled to be held in fully in parallel with the Loebner Prize also at the University of Reading. And as Kevin Warwick (Professor of Cybernetics at the University of Reading) recently hinted, "hosting the Loebner Prize is a great opportunity for the University of Reading. The competition is all about whether a machine can now pass the Turing Test, a significant milestone in Artificial Intelligence. I believe machines are getting extremely close - it would be tremendously exciting if such a world first occurred in the UK, in Reading University in 2008. This is a real possibility."

09:50 WELCOME (Dr. Mark Bishop, Goldsmiths)

SESSION 1: CHAIR: Prof. Kevin Warwick (Reading)
10:00 Baroness Susan Greenfield (Oxford)

11:00 ANNOUNCEMENT ON THE TURING TEST & COFFEE

SESSION 2: CHAIR: Prof. John Barnden (Birmingham)
11:20 Prof. Selmer Bringsjord (Rensselaer Polytechnic Institute, USA)
12:10 Dr. Michael Wheeler (Stirling)

13:00 BREAK FOR LUNCH

SESSION 3: CHAIR: Dr. Mark Bishop (Goldsmiths)
14:00 Dr. Andrew Hodges (Oxford)
14:50 Prof. Luciano Floridi (Herts/Oxford)

15:40 LOEBNER ANNOUNCEMENT & COFFEE

SESSION 4: CHAIR: Ms. Huma Shah (Reading)
16:00 Prof. Maggie Boden (Sussex)
16:50 Prof. Owen Holland (Essex)

17:40 PANEL DISCUSSION (ALL):
Is there a canonical Turing Test?

18:00 CLOSE (approx. timing)

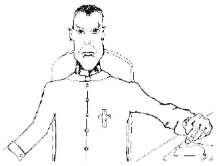
If Professor Warwick's suspicions are well grounded, this AISB symposium on the Turing Test promises to be a truly groundbreaking event and we anticipate considerable media attention around the 11am announcement. Nonetheless if you are interested in attending the event in person there are places available; these can be reserved by sending your name+address and cheque made payable to AISB for either: £25 (public); £20 (member, AISB) or £10 (full-time student/unemployed/OAP) to: Dr. J.M.Bishop, Dept. Computing, Goldsmiths, New Cross, London, SE14 6NW.

AISB Convention 2009

The 2009 AISB Convention will take place at Heriot-Watt University in Edinburgh from 6-9th April 2009. The theme of the convention will be Adaptive and Emergent Behaviour and Complex Systems.

ECAI 2012

The AISB, in collaboration with other societies and various universities, is planning to bid to host the European Conference on Artificial Intelligence in 2012. More details will be announced nearer the time.



Fr. Hacker's Guide for the Young AI Researcher

Cognitive Divinity Programme
Institute of Applied Epistemology

Research Excellence Framework Special

We interrupt the "Life of A. Hacker" to bring you this special and important supplement to Hacker's Guide. The main determinant of your score in the new Research Excellence Framework will be citation counts. Maximising citations to your work will be your passport to appointments, promotions, and research fame. This can be a blessing, not a curse, if you learn how to be cited

We have already covered the elementary citation increasing techniques. Guide 4.1 suggested publishing something obvious but faulty, so that everyone would cite it as a straw man. Guides 4.2 & 4.3 described techniques of coercion and alliance formation. Guides 4.4 & 4.5 discussed the judicious use of name changes to inherit the citations of others and to ensure that your self-citations would not be excluded from citation counts. Hacker Enterprises' CLUB (Collaboration of Lots of User Back-scratching; see Guide 7.6) extended cross-citation alliances to software. However, the latest methodological research has suggested several, more advanced, citation-increasing techniques, which it is now my urgent duty to propagate.

Surveys are the most cited form of research paper, so you should publish as many of these as possible. Fortunately, AI presents plenty of opportunities. Our field continually spawns new subfields, about which ambitious AI researchers are always anxious to learn. Moreover, these 'new' subfields are often closely related to earlier, now out-of-fashion subfields. So surveys of these earlier subfields can provide the essential ingredients of the new survey. The only problem is to spot the opportunity earlier than your rival surveyors. Here's where Hacker Enterprises' ABFAB (Aloysius's Bellwether for Fashions And Buzzwords) is invaluable. Its softbots will constantly search the internet for new survey opportunities for you.

The increasing publication in non-Latin alphabets, such as Arabic, Cyrillic, and Kanji, has made possible an almost undetectable self-citation technique. Cross-citing your own papers will increase your citation count, but at the cost of having to write lots of papers in many different alphabets. With the aid of Hacker's ALTER (Artificial Language Translation into Esoteric Renditions) you can automate this process. You provide an original research paper in English; ALTER will automatically translate this article into a wide variety of languages in non-Latin alphabets, each of which cross-cites all the others, including the English original. Furthermore, Hacker Enterprises' SOW (Submit to Outlets Worldwide) will not only find suitable foreign outlets for each of these translations, but will automatically handle the whole submission process for you in the appropriate language.

It is clear from numerous studies that authors don't read many of the papers they cite. You may, therefore, safely claim that a paper of yours anticipated one of the current fashionable research areas. To avoid offence, many authors will play safe and cite your pioneering work. If, following our advice, your paper is deep but impenetrable (see Guide 13), then such claims will be difficult to refute, even by those few who do read your paper. Better still, if you have written it in Cyrillic.