Evertsz, R., Ritter, F. E., Busetta, P., & Pedrotti, M. (2008). Realistic behaviour variation in a BDI-based cognitive architecture. In *Proceedings of SimTecT 2008*, http://www.siaa.asn.au/library_simtect_2007.html. SIAA Ltd.: Melbourne, Australia.

Realistic Behaviour Variation in a BDI-based Cognitive Architecture

Dr Rick Evertsz; Dr Frank E. Ritter; Mr Paolo Busetta; Mr Matteo Pedrotti
AOS Group, PSU (Frank Ritter)
[rick.evertsz, paolo.busetta; matteo.pedrotti]@aosgrp.com; frank.ritter@psu.edu

Abstract. BDI (Belief/Desire/Intention) models have a long and successful track record of use in CGFs (Computer-Generated Forces). Over the last two decades, cognitive architectures have been used to predict human performance in a wide range of tasks. A third stream of work seeks to develop architectures that account for behaviour moderation, by modelling affective and physiological factors. This paper describes the union of these three strands in the form of a novel, BDI-based, moderated cognitive architecture, CoJACK. CoJACK combines the ease-of-use of a graphical BDI language with the constraints of a cognitive architecture. Affective and physiological modelling is achieved by overlaying moderators on the cognitive architecture. CoJACK includes tracing and monitoring tools to aid with model development and debriefing of simulation runs. A situation awareness model is also provided.

1. INTRODUCTION

Simulation is essential in the modern military environment, however, in many current synthetic environments the SAF entities (Semi-Autonomous Forces) are completely consistent and will execute the same task in the same way each time. In the real world, this is not the case. The choice of strategies and the ordering of sub-strategies will vary across individuals and will vary for a given individual across time. When such variance is not included in a model, it makes adversaries, allies and neutral personnel too predictable because they will always do the same thing at the same time in the same way.

This paper describes CoJACK™, an agent-oriented, moderated cognitive architecture. CoJACK augments standard JACK® [1] agents with mechanisms including focus of attention, memory retention and recall, as well as timing and errors. CoJACK includes a moderator layer that affects the parameters of the cognitive architecture with factors including fear, morale, leadership and fatigue, amongst others.

Because CoJACK is implemented as a generic architecture, it can be added to standard JACK tactical models with minimal change to those models. This means that legacy JACK models can be updated to incorporate realistic human variation without re-implementation. An important feature is that CoJACK does not compromise JACK's ease-of-use and representational clarity. JACK represents and executes tactics in a manner that maps well to Subject Matter Experts' (SMEs) introspections about their own reasoning processes.

Additionally, CoJACK includes an SA (Situation Awareness) model and provides a suite of tracing and monitoring tools to support the construction of moderated cognitive models.

2. THE NEED FOR BEHAVIOUR VARIATION

The representation of human behaviour in simulation has a long history, beginning with the use of procedural programming languages such as FORTRAN. Such models were difficult to modify or extend, and this led to the development of scripting capabilities in SAFs. However, scripted behaviour models tended to lack flexibility and were not well suited to the representation of dynamic human decision-making capabilities.

These drawbacks led to the use of Artificial Intelligence (AI) languages such as Soar [2] and JACK. Although this resulted in convincing high-level decision making, variation had to be explicitly programmed if it was provided at all.

Other research endeavours have focused on sources of variability that derive from physiology and emotion. For example, PSI integrates cognition, emotion and motivational factors [3, 4]. PMFServ [5] is an architecture that also represents affective states.

Emotion, affect, motivation, individual differences, behaviour moderators and other sources of human variability are increasingly being seen as factors that can and often do influence cognition and performance in important ways in synthetic environments. A major historical emphasis in military simulation has been on training and analysis. This has led to a strong focus on repeatability of results, which was seen as being at odds with a desire for variance in performance. In the past, variance was also intentionally suppressed in simulations because it was thought that variance in real behaviour could be controlled through doctrine and training. However, while this may reduce variance it will never completely eliminate it, and further reflection and reading suggests that it remains. Grossman [6] notes that combat participation varies due to several factors, and a recent simulation using CoJACK [7] showed that fear could be used to influence commanders' choices between legal actions based upon a fixed set of rules of engagement.

Thus, although many applications of simulation are focused on obtaining predictable and repeatable behaviour, this can lead to a false sense of confidence in the veracity of the outcome, and therefore a less useful simulation of human performance. We argue that it is important to represent the inherent variability in human behaviour in a principled manner. Only then is it possible to predict *potential* for variation, as well as the *range* of that variation. By varying the initialisation parameters, multiple runs produce the full range of behaviours that are implicit in the models. Furthermore, if data is available on the distribution of parameter values in the human population, our approach can associate probability values with each outcome.

2.1 BDI Agents in Behaviour Modelling

BDI (Belief/Desire/Intention) agents have a relatively long history in CGFs (Computer Generated Forces), e.g. SWARMM [8]. The BDI paradigm is well suited to CGFs because, like humans, BDI agents can be reactive or proactive, depending on the demands of the situation. When executing a course of action, a BDI agent can always switch its attention to a significant event in its environment.

The JACK BDI lineage in particular is distinguished by its intuitive, SME-friendly representation language. This lineage began with PRS (Procedural Reasoning System) which was developed for fault diagnosis of Space Shuttle missions [9], transitioned through dMARS [10], and is currently represented by JACK. All three languages offer a graphical representation of "plans" (the BDI representation of the reasoning process). This graphical plan representation was key to the choice of JACK as the foundation for CoJACK. State of the art cognitive architectures, with their fine-grained production rule structure, can make models tedious to build and difficult for SMEs to understand (cf. ACT-R [11]).

2.2 Brief Overview of JACK

JACK is an agent-based programming language with the following major constructs:

- Events – are the central motivating factor in agents. Events are generated in response to external stimuli or as a result of internal agent computation.
- Plans – are procedures that define how to respond to events. When an event is generated, JACK computes the set of plans that are applicable to the event. These plans are subjected to a process of deliberation, where the agent selects the plan that will form its next intention. Plans have a body that defines the steps to be executed in response to the event. Non-deterministic choice allows the agent to try alternative plans to achieve the goal.
- Beliefsets – are used to represent the agent's declarative beliefs in a first order, tuple-based relational form. Beliefsets are analogous to the Working Memory (WM) of a production system.
- Intentions – are the currently active plan instantiations, i.e. the plan instances that the agent is committed to. Plans are abstract entities that describe potential patterns of thought and action. A plan becomes an intention when the agent instantiates it with symbolic references to concrete entities in the environment, and commits to its execution.

2.3 What is a Cognitive Architecture?

CoJACK is a BDI-based cognitive architecture. A cognitive architecture is a computational system that models the structural properties of the human cognitive system — the information processing mechanisms that are fixed across tasks. As such, it constrains the models that can be implemented therein by only allowing the definition of models that fit within its structural boundaries. This provides a significant advantage over ad-hoc approaches to modelling human behaviour: as a theory of cognition, cognitive architectures provide a more principled and testable framework for investigating the mechanisms of human information processing.

Because a cognitive architecture is concerned with structural properties, it defines the invariant aspects of the human cognitive system. A number of cognitive architectures are widely used in the cognitive modelling community; most notably, Soar [12] and ACT-R [11]. Nevertheless, their history is relatively short and cognitive psychologists are only at the beginning of the road to a unified theory of cognition that is consistent with the large body of psychological data.

3. CoJACK

Learning from the deficiencies of PRS and dMARS, the JACK development team (who worked with the designer of PRS and subsequently developed dMARS), designed JACK to be lightweight and efficient. This was to enable its use in resource-limited computing environments (e.g. unmanned aircraft systems). With this in mind, CoJACK was designed to be an extension that does not sacrifice JACK's efficiency in cases where the CoJACK extensions are not used.

CoJACK's cognitive architecture can be disabled at compile time, resulting in behaviour that is the same as that produced by normal JACK agents. As a cognitive architecture, CoJACK extends JACK so that it simulates the time taken to process information, and it adds forgetting and error-prone retrieval of declarative and procedural knowledge. CoJACK augments this with a moderator layer that allows the cognitive architecture's performance to be modulated by physiological factors (e.g. level of caffeine in blood) as well as affective dimensions (e.g. aggression).

The cognitive architecture and moderator layer represent the core functionality of CoJACK (see figure 1). To support the model builder and the analyst, CoJACK includes a suite of graphical tracing an monitoring tools, as well as an implementaton of Endsley's model of situation awareness [13].

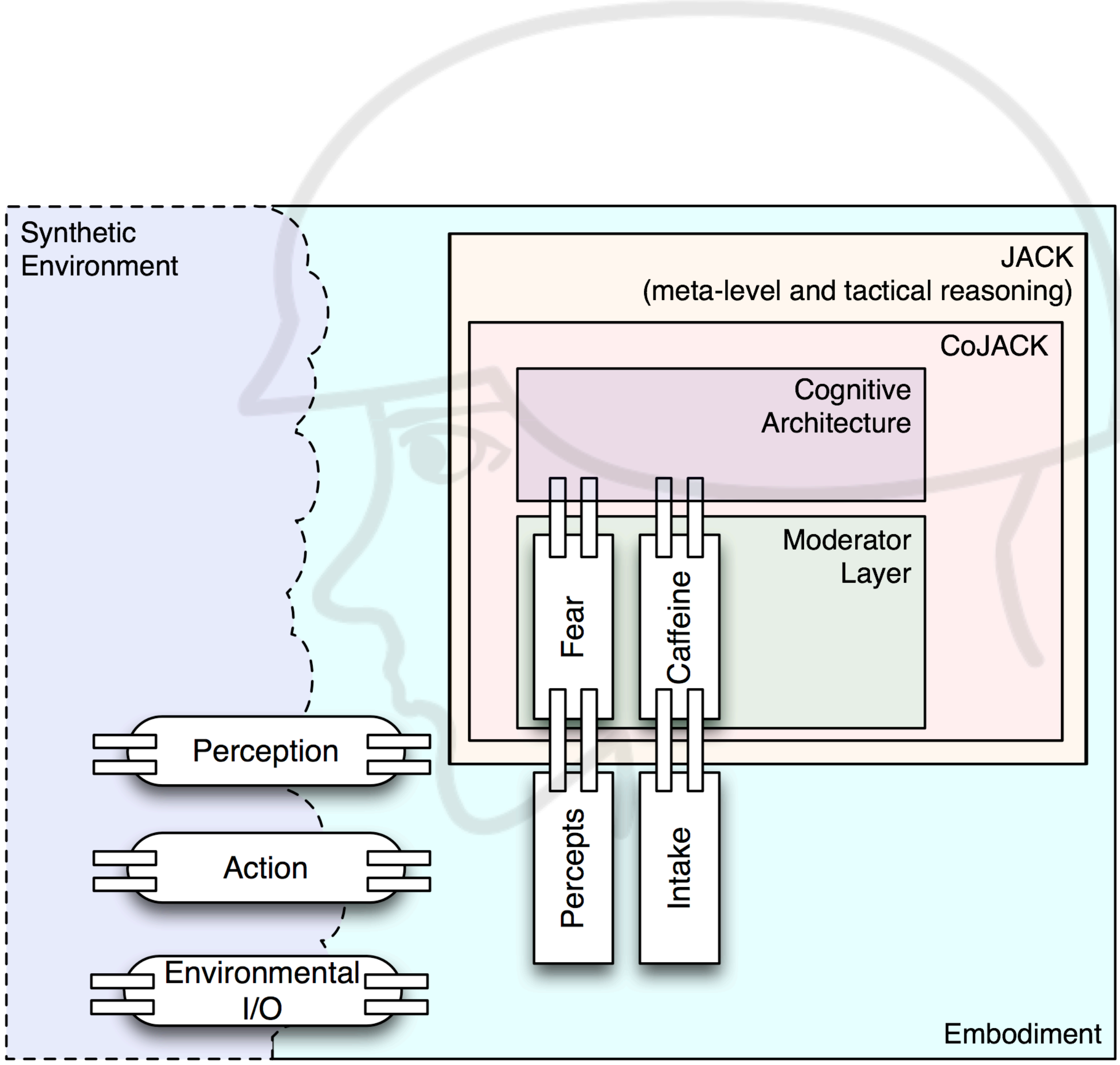

Figure 1: CoJACK architecture in the context of a synthetic environment

3.1 High-Level Architecture of a CoJACK Agent

This section presents a conceptual model of how a CoJACK agent fits in with JACK and a typical SE, as illustrated in figure 1.

- A minimal CoJACK agent has an embodiment that represents its sensory and motor capabilities, but can also include a model of its body. Typically, this embodiment is largely in the SE, with some overlap with the agent itself (as represented by the dotted boundary in figure 1). Perception and action are mediated by the interface between the agent and the SE, and this interface can also include environmental information such as temperature. The percepts and incoming environmental data can be routed to moderators in the agent (in figure 1, percepts form an input to the Fear moderator, and caffeine uptake is routed to the Caffeine moderator).
- The moderators plug into the cognitive architecture and modulate the values of the architecture's parameters. Cognitive parameters are described in the next section. The moderators themselves can have internal state in the form of one or more reservoirs that have associated decay functions. For example, the caffeine moderator has a single reservoir that represents the amount of caffeine in the bloodstream. The reservoir level gradually decreases according to a decay function that represents the normal rate of excretion of caffeine.
- The JACK level is responsible for tactical reasoning and can also perform meta-level reasoning (e.g. thinking about its decision making). When CoJACK is enabled, the JACK level is subject to the constraints of the cognitive architecture (e.g. timing of reasoning steps, error-prone retrieval from memory, etc.).

3.2 Details of the CoJACK Core

The goal for CoJACK was to develop a cognitive architecture that predicts the timings and errors in human performance without compromising the usability of the BDI representation. We start with the assumption that humans share a common cognitive architecture and physiology, and that variation results from individual differences in knowledge and the values of the architecture's and physiology's parameters.

CoJACK addresses the following major aspects of cognition:

1. **Cognition takes time.** In CGFs, reasoning steps execute effectively instantaneously, as does memory retrieval and update. If, for example, the granularity of the simulation is 1s, a model can run through a series of decision-making steps in 1s (simulated or real) that would take a human half an hour. In CoJACK, reasoning steps have a default duration, but this can be moderated (e.g. by fatigue).
2. **Potential for failure to retrieve a belief (declarative knowledge).** Working Memory (WM) is the mechanism of human cognition that maintains information during processing. It is limited in capacity as evidenced by decreased performance in tasks that require many temporary items to be held in memory. In contrast, CGFs tend not to reflect WM limitations. CoJACK imposes WM limits on memory retrieval – a belief will not be retrieved if its activation level falls below the retrieval threshold.
3. **Possible failure to remember the next step in a course of action (procedural knowledge).** In a similar manner to beliefs, intentions have an activation level in CoJACK. If this is below the retrieval threshold, the agent will fail to follow the next step of its current intention.
4. **Retrieval of the wrong belief.** CoJACK supports fuzzy retrieval of beliefs.
5. **Limited focus of attention.** WM, as it relates to the agents active intentions, ensures that the agent can only focus on a limited number of activities at once.
6. **Cognition can be moderated.** Moderators can dynamically alter the parameters of the cognitive architecture, thereby affecting the agent's processing.

The cognitive architecture alters the symbolic reasoning of the agent by affecting it at the sub-symbolic level. This borrows from the approach taken in ACT-R [11]. Further details are available [14].

3.3 Moderator Representation

CoJACK moderators are characterised in terms of time-based functions that are applied to controllable parameters (primarily cognitive parameters, but potentially other user-defined parameters).

Each agent's moderators are initialised at simulation start up, defining parameter base values, the moderator functions and prior moderator exposures (e.g. an agent's initial level of fatigue).

Moderators are represented as a network of interconnected function elements that state how moderator receptors are linked to parameter variation. Thus, a moderator is a group of mathematical functions that specifies how a collection of parameters varies over time with respect to exposure events. Moderator functions can maintain internal state (e.g. a caffeine reservoir).

4. SITUATION AWARENESS IN CoJACK

SA is defined [15] as:

"… the ability to have accurate real-time information of friendly, enemy, neutral, and non-combatant locations; a common, relevant picture of the battlefield scaled to specific levels of interest and special needs…" (Glossary-7)

and in a more general manner by Endsley [13] as:

"… the perception of elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." (p. 97)

At all three of Endsley's levels (perception, comprehension, and projection), SA can be affected by a wide range of psychological factors, and ultimately determines the timeliness, appropriateness and effectiveness of decisions and actions. It is fundamental to human performance – decisions that are based on incomplete or incorrect knowledge of the situation are less likely to yield optimal action and may be counter-productive. Hartel et al. [16] report that SA was the leading causal factor in the 175 military aviation mishaps reviewed in their study. Hence, SA is seen as central to effective decision-making, and is now the subject of considerable research in the fields of civil aviation, road safety, counter-terrorism and military science.

Situation awareness forms the input to the decision-making process and is central to the behaviour of agents (whether human or synthetic), particularly those that interact with an unpredictable environment. Variations in SA are believed to account for a considerable amount of the variability in human behaviour in military scenarios. As such, it will be important to be able to observe the evolution of the SA of a given agent in a simulation. It will be of particular interest to note when and where an agent's SA is incorrect or incomplete. With this in mind, CoJACK provides explicit support for SA modelling.

Ultimately, the contents of the SA model must be open to inspection by the analyst. Based on a study of military and scientific experts looking at a behaviour model [17, 18, 19], it was found that they want:

- to know what the structure of the model is,
- to know how to use and inspect the model, and
- to know why the model does what it does.

The first and third questions are related to SA – how the model represents its environment and what it has concluded about the current and future state of its world. This information can be provided in a graphical and/or textual form. A toolkit like VISTA [20] provides some of the required functionality.

4.1 The Endsley Model of SA

The Endsley [13] definition of SA is probably the most widely accepted. It is domain-independent and is stated in terms of process. Nevertheless, the definition implicitly addresses the products of situation assessment (i.e. percepts, meaning and future status).

The Endsley model consists of three levels: (1) perception, (2) comprehension and (3) projection. We term these **L1-SA**, **L2-SA** and **L2-SA** respectively We have implemented the Endsley model in CoJACK, based on a computational interpretation. Perception (L1-SA) is the first step of situation assessment and involves the acquisition of environmental data that is relevant to the task at hand. Comprehension (L2-SA), is concerned with further processing of L1-SA, inferring its implications. This inference process is primarily deductive but can also be abductive. Projection, (L3-SA) is the most complex aspect of situation awareness. It uses the results of L1-SA and L2-SA to predict what will happen in the near term. In an adversarial situation, projecting into the future might involve reasoning about the goals and actions of the opponent and mentally modelling how the future could play out.

The Endsley model is qualitative in nature and the classification of knowledge into the three levels is sometimes subjective. For some people, discerning a given attribute of the environment may involve inference (i.e. L1-SA), but for others with more experience, that attribute will be perceived without the need for explicit reasoning (e.g. the way a chess master perceives a board position).

In CoJACK, the three levels of SA are represented by transparently annotating beliefs with their SA level. The annotations are transparent in the sense that a plan that references a given belief, works whether or not the belief is annotated with SA-level information

- Beliefs that derive directly from elements in Sensory Memory are automatically tagged as L1-SA.
- Beliefs that are created by plans, from L1-SA and/or L2-SA are automatically labelled as L2-SA.
- Annotating L3-SA beliefs is not so straightforward to automate and requires that the model builder explicitly annotate L3-SA plans.

5. TRACING AND MONITORING TOOLS

CoJACK comes with a GUI for defining moderators and testing/monitoring them in advance of their inclusion in an agent model.

CoJACK is instrumented with three main types of reporting tool:

1. System event logging (e.g. start up of CoJACK, notification of configuration data, error conditions), aimed at post-mortem analysis of runs in a production environment.
2. Detailed execution tracing used for on-line monitoring of CoJACK. This is mainly used for debugging in a development environment.
3. Statistical data reporting, aimed at model analysis both in a production and development environment.

Furthermore, a GUI is available for on-the-fly changes to cognitive parameter values, moderator formulae and to inject moderator events during execution.

6. SUMMARY

We have described a novel, BDI-based, moderated cognitive architecture, CoJACK. CoJACK combines the ease-of-use of a graphical BDI language with the constraints of a cognitive architecture. Affective and physiological modelling is achieved by overlaying moderators on the cognitive architecture. CoJACK includes tracing and monitoring tools to aid with model development and debriefing of simulation runs. The Endsley model of SA has been implemented and can be used to investigate how SA impacts upon behaviour.

ACKNOWLEDGEMENTS

This work was supported by the UK MOD's Analysis, Experimentation and Simulation corporate research programme (Project No: RT/COM/3/006). We thank Colin Sheppard for useful discussions regarding all aspects of this work.

REFERENCES

1. Busetta, P., Rönnquist, R., Hodgson, A. & Lucas, A., (1999) "JACK intelligent agents - Components for intelligent agents in JAVA" *AgentLink News Letter*, vol. 2, pp. www.agent-software.com/white-paper.pdf.
2. Jones, R. M., Laird, J. E., Nielsen, P. E., Coulter, K. J., *et al.*, (1999) "Automated intelligent pilots for combat flight simulation" *AI Magazine*, vol. 20, pp. 27-41.
3. Bartl, C. & Dörner, D., (1998) "PSI: A theory of the integration of cognition, emotion and motivation" *Proceedings of the 2nd European Conference on Cognitive Modelling*, pp. 66-73, Thrumpton, Nottingham, UK, Nottingham University Press.
4. Bach, J. (2007) *Principles of synthetic intelligence: Building blocks for an architecture of motivated cogntion,* , Oxford University Press, New York, NY.
5. Silverman, B. G., (2004) "Human performance simulation" in J. W. Ness, D. R. Ritzer, V. Tepe (Eds.), Elsevier, Amsterdam, pp. 469-498.
6. Grossman, D. (1996) *On Killing,* , Back Bay Books, Little Brown and Company, New York, NY.
7. Evertsz, R., Ritter, F. E., Russell, S. & Shepherdson, D., (2007) "Modeling Rules of Engagement in Computer Generated Forces" *Proceedings of the 16th Conference on Behavior Representation in Modeling and Simulation*, pp. 123-34, Orlando, FL, U. of Central Florida.
8. Murray, G., Steuart, D., Appla, D., McIlroy, D., *et al.*, (1995) "*The challenge of whole air mission modelling*" *Proceedings of the Australian Joint Conference on Artificial Intelligence*, Melbourne, Australia.

9. Georgeff, M. P. & Ingrand, F. F., (1989) "Monitoring and control of spacecraft systems using procedural reasoning" *Proceedings of the Space Operations Automation and Robotics Workshop*.

10. d'Inverno, M., Kinny, D., Luck, M. & Wooldridge, M., (1998) "A formal specification of dMARS" *Proceedings of the fourth International Workshop on Agent Theories, Architectures, and Languages (ATAL)*, pp. 155-76.

11. Anderson, J.R. (2007) *How can the human mind exist in the physical universe?,* , Oxford University Press, New York, NY.

12. Lewis, R. L., Huffman, S. B., John, B. E., Laird, J. E., *et al*., (1990) "Twelfth Annual Conference of the Cognitive Science Society" *Soar as a Unified Theory of Cognition: Spring 1990*, pp. 1035-42, Cambridge, MA.

13. Endsley, M. R., (1988) "Design and evaluation for situational awareness enhancement" *Proceedings of the Human Factors Society 32nd Annual Meeting*, pp. 97-101, Santa Monica, CA, Human Factors Society.

14. Evertsz, R., Ritter, F. E., Busetta, P., Pedrotti, M. & Bittner, J. L., (2008) "CoJACK – Achieving Principled Behaviour Variation in a Moderated Cognitive Architecture" *Proceedings of the 17th Conference on Behavior Representation in Modeling and Simulation*, Orlando, FL, U. of Central Florida.

15. TRADOC (1998) *Land combat in the 21st Century (draft),* , U. S. Army Training and Doctrine Command, Fort Monroe, VA: Headquarters.

16. Hartel, Smith, K. & Prince, C., (1991) "Defining Aircrew Coordination: Searching Mishaps for Meaning" *Proceedings of the Sixth International Symposium of Aviation Psychology*, Columbus, Ohio, Ohio State University.

17. Avraamides, M. & Ritter, F. E., (2002) "Using multidisciplinary expert evaluations to test and improve cognitive model interfaces" *Proceedings of the 11th Computer Generated Forces Conference*, pp. 553-62, Orlando, FL, U. of Central Florida.

18. Councill, I. G., Haynes, S. R. & Ritter, F. E., (2003) "Explaining Soar: Analysis of existing tools and user information requirements" *Proceedings of the Fifth International Conference on Cognitive Modeling*, pp. 63-8, Bamberg, Germany.

19. Councill, I. G., Haynes, S. R. & Ritter, F. E., (2003) "Explaining Soar: Analysis of existing tools and user information requirements" *Proceedings of the Fifth International Conference on Cognitive Modeling*, pp. 63-8, Bamberg, Germany, Universitäts-Verlag Bamberg.

20. Taylor, G., Jones, R. M., Goldstein, M. & Frederiksen, R., (2002) "VISTA: A generic toolkit for visualizing agent behavior" *Proceedings of the 11th Computer Generated Forces Conference*, pp. 29-40, Orlando, FL, U. of Central Florida.