The University of Memphis — educational institution
Mathematical Sciences Department &
Institute for Intelligent Systems
The University of Memphis
Memphis, TN 38152

Conceptual Intelligent Distribution Agents

> "This proposal is submitted pursuant to the ONR Broad Agency Announcement (BAA) # 97-024 for Long Range Scientific Projects, published in the Commerce Business Daily, issue No. PSA-1634 of 17 July 1997."

Funding requested — $67,495

Duration of effort — Fiscal Year 1998

Stan Franklin — Principal Investigator
(901) 678-3142
fax: (901) 678-2480
stan.franklin@memphis.edu
http://www.msci.memphis.edu/~franklin

Sally Johnson — Business Contact
(901)-678-4146
(901)-678-2199
sjohnson@cc.memphis.edu

Abstract

Like the Roman god Janus, this project has two faces, its science face and its engineering face. Its science side will flesh out the global workspace theory of consciousness, while its engineering side explores architectural designs for information agents that promise more flexible, more human-like intelligence within their domains. The architectures and mechanisms that underlie consciousness and intelligence in humans can be expected to yield information agents that learn continuously, that adapt readily to dynamic environments, and that behave flexibly and intelligently when faced with novel and unexpected situations. The basic idea is to achieve more human-like intelligence by doing it like humans do it. This technology should produce "intelligent active information agents ... that [can] automate and improve how people integrate and analyze information, make decisions under uncertainty and communicate knowledge and decisions to multiple sites. "

Built upon and extending previously developed clerical agent architectures, this project is intended as a conceptual study of the global workspace approach to autonomous agent design. We will produce a conceptual design for an intelligent information agent who, with minimal input via email in natural language, will seek out and coordinate information from various sources, and will return detailed suggestions for actions based on the information gathered. 'Living' in a unix system, this agent will function with no human supervision, and will perform a task currently thought of as only possible for humans.

The particular domain chosen for this proof-of-concept project is that of a conceptual intelligent distribution agent that would implement a market-driven job assignment system. . Such an assignment system would empower both individuals and commands to find mutually beneficial job assignments. This domain is sufficiently challenging in that it requires information retrieval from both local databases, and remote sources, inferences from this information, and continual learning about both its sources and its users. The designs developed for this project will be applicable to many other domains of concern to the Navy. Here we propose the conceptual design for such an agent, not its implementation.

Relation to the objectives of ONR and scientific issues concerning Manpower and Personnel.

The Navy must compete in an ever-changing marketplace to attract and retain qualified personnel. The Navy must also maintain a sufficient levelof readiness in a world where its mission may be rapidly changing. The ability to maintain personnel readiness by placing the right person in the right job at the right time is far more difficult today due to the increaseduncertainty surrounding potential Navy roles of the future. An intelligent, flexible, creative distribution system would be most useful in responding to the needs of a new, dynamic Navy. The core of such an endeavor would be a market-driven job assignment system which would empower both individuals and commands to find mutually beneficial job assignments.

Such a flexible, intelligent, distribution system would provide a perfect proof-of -concept domain for conscious software, as well as producing an important new advanced technology for the use of the Navy. In particular, we propose to design Intelligent Manpower, Personnel and Training (MPT) Agents that would reside at localcommands (e.g., ships) and BuPers that provide automatic management and expert assistance on MPT matters. These agents when implemented and trained could enable collaborative detailing, training, and career counseling, provide expert advice, and reduce administrative manpower requirements.

Currently the distribution system attempts to balance the preferences of the Sailor with the needs of the ships in the fleets and with the priorities of the Bureau of Personnel, which considers the Sailor, the fleets, costs, and other factors. When the time comes for a Sailor to change jobs, the decision is made based on all of the factors in effect at that time. Conscious software agents could be constantly looking for matches in such a way that all three parties' needs are satisfied better than with the current system. Here we propose to design such agents.

Technical Summary

Artificial intelligence pursues the twin goals of understanding human intelligence and of producing intelligent software. Designing, implementing and experimenting with autonomous agents (Franklin and Graesser 1997) furthers both these goals in a synergistic way. Each such design, including architecture and mechanisms, gives rise to a theory that says humans operate according to that design. Conversely, every theory of human cognition constrains the design of autonomous agents built to implement the theory. Changes in either the theory or the design often require changes in the other. Thus what we learn about either enhances the other. (See Franklin, 1997, for a expanded account of this argument.).

Here we propose to design, implement and experiment with an autonomous internal travel clerk constrained by Baars' global workspace theory of consciousness (Baars 1988, 1997), to be described briefly below. The enterprise proposed here stands upon earlier work on Virtual Mattie, an intelligent clerical agent (Franklin et al 1996). At this writing, the VMattie design is complete and the coding is also complet except for the final integration of modules. These modules are running correctly individually. Experimental study should begin before September 1, 1997.

The VMattie design implements a major chunk of the global workspace theory. An extension of this design, called 'Conscious' Mattie, will implement the entire global workspace theory. This work is also well along, and is proceeding in parallel with the development of VMattie. Coding of CMattie should be complete and experimentation begun well before funding for this travel agent project is expected (September 1, 1998). The research team for all three projects consists of two faculty members and seven doctoral students, not all of whom would be supported by this grant.

The global workspace theory (Baars 1988, 1997) postulates that human cognition is implemented by a multitude of relatively small, special-purpose processes, almost always unconscious. (Cognition is implemented by a multiagent system.) Coalitions of such processes find their way into a global workspace (and into consciousness). This limited capacity work space serves to broadcast the message of the coalition to all the unconscious processors, in order to recruit other processors to join in handling the current novel situation, or in solving the current problem. All this takes place under the auspices of contexts: goal contexts, perceptual contexts, conceptual contexts, and/or cultural contexts. Each context is itself a coalition of processes. There's much more to the theory, including attention, learning, action selection, problem solving, etc.

Though quite comprehensive, Baars' theory is, as a psychological theory should be, quite abstract, offering general principles and broad architectural sketches. Questions of architectural detail, that is of just how functional components fit together and who talks to whom, are left open. Also open are questions of mechanisms, that is of how these components do what they are claimed to do. For example, in Baars' presentation the various types of contexts (perceptual, conceptual, goal contexts) are lumped architecturally. Though distinguished functionally, their architectural relationships, as well as their mechanisms, are left unspecified. As a good theory should, this one raises as many questions as it answers.

The long term scientific objective of this project is to propose answers to many of these questions by fleshing out the theory with a detailed architecture and specified mechanisms. Our proposed answers will then be tested in computer models of the theory. The computer models will tell us if the proposed architecture and mechanisms work computationally, and if they are consistent.

Turning to the engineering side, artificial intelligence has often been taken to task for producing software systems that are 1) ungrounded conceptually and thus dependent on human users for their input and for the interpretation of their output, 2) brittle at the edges, that is, given to complete breakdown or, worse, nonsense output when taken slightly out of their domain, and 3) unable to successfully cope with the frame problem (Agre and Chapman 1988, Brooks 1990a, Brown 1987, Cliff 1991, Crocket 1994, Dreyfus and Dreyfus 1987, Edelman 1987). Though classical AI's position on each of these issues is defensible, there is also an element of truth in each.

One suggestion is that AI should switch from making simple models of complex systems, and hoping that they scale up, to making complex models of simple systems, say insect-like systems (Brooks 1990a, Cliff 1991). Many such systems have been built (Beer 1990, Brooks 1990b, Johnson and Scanlon 1987, Maes 1991b, Pfeifer 1996), both in hardware and in software. Though successful in their own terms, none of them promise to scale up to software systems with flexible, adaptive, human-like intelligence, even in a narrow domain.

A second suggestion, closely related to the previous, recommends that AI concentrate on autonomous software agents (Franklin and Graesser 1997), thereby eliminating the grounding problem and the frame problem (Maes 1991a, Neisser 1993). Many such agents have been built (Etzioni and Weld 1994, Ferguson 1995, Hayes-Roth 1995, Maes 1994, Song, Franklin and Negatu 1996). Again, though successful in their own terms, these systems still suffer from brittleness, and show little promise of scaling up to flexible, adaptive systems with human-like intelligence within and at the boundaries of their domains.

The long term engineering goal of this project is to discover principles enabling the development of flexible, adaptive software systems that display human-like intelligence within narrow, but not too narrow, domains. To accomplish this, we will design architectures and mechanisms constrained by the global workspace theory, implement these designs as autonomous software agents, and experiment with these agents. The evaluation of these experiments will suggest modifications of the architecture and mechanisms, perhaps still within the constraints of global workspace theory. Though, the time scale of this proposal will allow only for a conceptual design, we envision seeking further support for several iterations of cycle described above, greatly increasing the likelihood of our finding the enabling principles we are seeking. In this project the science face and the engineering face work together synergisticly, as described above. The basic idea is to achieve more human-like intelligence by doing it like humans do it.

Our first software model of global workspace theory will be an extension the architecture and mechanisms of an autonomous software agent, VMattie, currently under development (Franklin et al 1996). VMattie is designed to compose and email out weekly seminar announcements, having communicated by email with seminar organizers and announcement recipients in natural language, all with no human involvement other than email messages. Her architecture, as currently designed, implements a major piece of global workspace theory. This is not surprising since the VMattie architecture was constrained by the action selection paradigm of mind as espoused in Artificial Minds (Franklin 1995). This paradigm and global workspace theory are quite compatible. VMattie's architecture is based on the fusion of extended versions of behavior nets (Maes 1990) and the Copycat architecture (Hofstadter and Mitchell 1994). VMattie's behaviors correspond to the goal contexts of global workspace theory, while the Copycat codelets play the role of processes. VMattie is up and running and in the debugging stage at this writing (September 25, 1997). Experimentation and evaluation should be complete by the end of October 1997.

We will then turn to implementing extensions of the VMattie architecture to complete the modeling of the global workspace theory. (The design of these extensions is almost complete and proceeding in parallel with the implementation and evaluation of VMattie.) Initially both pandemonium theory (Jackson 1987) and sparse distributed memory (Kanerva 1988) will be used. Kanerva's mechanism will implement episodic memory, while Jackson's demons (codelets) will implement associative learning. During this same period, the limited capacity global workspace with broadcasting capabilities will be added again using pandemonium theory together with blackboard-like constructions and a fuzzy logic controller. We expect to have CMattie fully implemented and evaluated by July 1, 1998. (Our estimates have been amazingly accurate thus far.)

In the final phase of this an agent fully modeling the global workspace theory in a more challenging domain than that of VMattie will be designed and implemented. The conceptual design of this type of agent is being proposed here. Experimentation with these agents, their architectures and mechanisms, should yield the general architectural principles we seek. This agent should provide a proof of concept of "conscious" software agents in that it should perform a task heretofore only expected only of a human.

We've chosen conceptual intelligent distribution agents for the Navy as our proof of concept domain, because it provides a sufficiently challenging, but doable, domain. The architectures and mechanisms used will be derived from those of CMattie with extensions as needed. The design of these agents will proceed in parallel with the implementation and evaluation of CMattie.

The Navy must compete in an ever-changing marketplace to attract and retain qualified personnel. The Navy must also maintain a sufficient levelof readiness in a world where its mission may be rapidly changing. The ability to maintain personnel readiness by placing the right person in the right job at the right time is far more difficult today due to the increaseduncertainty surrounding potential Navy roles of the future. An intelligent, flexible, creative distribution system would be most useful in responding to the needs of a new, dynamic Navy. The core of such an endeavor would be a market-driven job assignment system which would empower both individuals and commands to find mutually beneficial job assignments.

Such a flexible, intelligent, distribution system would provide a perfect proof-of -concept domain for conscious software, as well as producing an important new advanced technology for the use of the Navy. In particular, we propose to design Intelligent Manpower, Personnel and Training (MPT) Agents that would reside at localcommands (e.g., ships) and BuPers that provide automatic management and expert assistance on MPT matters. These agents when implemented and trained could enable collaborative detailing, training, and career counseling, provide expert advice, and reduce administrative manpower requirements.

Currently the distribution system attempts to balance the preferences of the Sailor with the needs of the ships in the fleets and with the priorities of the Bureau of Personnel, which considers the Sailor, the fleets, costs, and other factors. When the time comes for a Sailor to change jobs, the decision is made based on all of the factors in effect at that time. Conscious software agents could be constantly looking for matches in such a way that all three parties' needs are satisfied better than with the current system. Here we propose to design such agents.

Personel

The research group currently working on these projects consists of two faculty members, Stan Franklin and Arthur C. Graesser, and seven doctoral students. Only Franklin and two doctoral students are to receive support under this proposal. The others will continue to work on the projects with support from elsewhere.

The PI, Stan Franklin, in addition to recent theoretical and practical work on autonomous software agents (Franklin 1997, Franklin, and Graesser 1997, Franklin et al 1996, Song, Franklin and Negatu 1996, Zhou and Franklin 1994), has recently reviewed the literature on intelligent software agents and cognitive agents (Franklin 1995, Franklin and Graesser 1997). Franklin also brings a strong background in neural nets (Franklin and Garzon 1996) and some research in genetic algorithms (Kilman, Dickerson, Franklin and Wong 1992) to the project.

References

Agre, Philip E. and Chapman, David (1988), "Indexicality and the binding problem," Proceedings of the AAAI Symposium on Parallel Models, 1-9.

Baars, Bernard J. (1988), A Cognitive Theory of Consciousness, Cambridge: Cambridge University Press.

Baars, Bernard J. (1997), In the Theater of Consciousness, Oxford: Oxford University Press.

Beer, Randall D. (1990), Intelligence as Adaptive Behavior, Boston: Academic Press.

Brooks, Rodney A. (1990a), "Elephants Don't Play Chess," in Pattie Maes editor, Designing Autonomous Agents, Cambridge MA: MIT Press.

Brooks, Rodney A. (1990b), "A Robot that Walks; Emergent Behaviors from a Carefully Evolved Network," P. H. Winston, ed., Artificial Intelligence at MIT, vol. 2, Cambridge MA: MIT Press.

Brown, Frank, M. (1987) ed., The Frame Problem in Artificial Intelligence, Los Altos, CA: Morgan Kaufman.

Cliff, D.T. (1991), "Computational Neuroethology: A Provisional Manifesto," in J.-A. Meyer and S. W. Wilson, editors, From Animals to Animats, Cambridge MA: The MIT Press, 29-39.

Crockett, L. (1994) The Turing Test and the Frame Problem: AI's Mistaken

Understanding of intelligence. Norwood, NJ: Ablex.
Dreyfus, Hubert L. and Stuart E. Dreyfus (1987), "How to Stop Worrying about the Frame Problem even though It's Computationally Insoluble," in Z. W. Pylyshyn, ed., The Robot's Dilemma, Norwood NJ: Ablex.
Edelman, Gerald M. (1987), Neural Darwinism: The Theory of Neuronal Group Selection, New York: Basic Books.
Etzioni, Oren and Weld, Daniel (1994). "A Softbot-Based Interface to the Internet." Communications of the ACM, (37) 7: 72-79.
Ferguson, I. A. (1995). "On the role of DBI modeling for integrated control and coordinated behavior in autonomous agents." Applied Artificial Intelligence, 9(4).
Franklin, Stan (1995). Artificial Minds. Cambridge, MA: MIT Press.
Franklin, Stan (1997). Autonomous Agents as Embodied AI,Cybernetics and Systems' Special issue on Epistemological Aspects of Embodied AI, 28 499-520. Available online at: http://www.msci.memphis.edu/~franklin/cyber.html/
Franklin, Stan and Garzon, Max (1996) "Computability via Discrete Neural Nets," in Rumelhart, Smolensky and Moser eds. Mathematical Perspectives on Neural Networks, Erlbaum, 1996, 41-84.
Franklin, Stan and Graesser, Art (1997) "Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents," Intelligent Systems III, Berlin: Springer Verlag, 21-35.
Franklin, Stan et al (1996), "Virtual Mattie-an Intelligent Clerical Agent" AAAI Symposium on Embodied Cognition and Action, Cambridge MA: AAAI Press.
Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. T. Spence, J. M. Darley, and D. J. Foss (Eds.), Annual Review of Psychology, Vol. 48. Palo Alto, CA: Annual Reviews Inc.
Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review 101, 371-95
Hayes-Roth, B. (1995). "An architecture for adaptive intelligent systems." Artificial Intelligence, 72 329-365.
Hofstadter, D. R. and Mitchell, M. (1994), "The Copycat Project: A model of mental fluidity and analogy-making." In Holyoak, K.J. & Barnden, J.A. (Eds.) Advances in connectionist and neural computation theory, Vol. 2: Analogical connections. Norwood, N.J.: Ablex.
Jackson, John V. (1987). "Idea For A Mind." SIGART Newsletter, July No. 101 23-26
Johnson, M. and Scanlon, R. (1987), "Experiences with a Feeling-Thinking Machine," Proceedings of the IEEE First International Conference on Neural Networks, San Diego, 71-77.
Kanerva, P. (1988). Sparse Distributed Memory. MIT Press.
Kilman, D. Dickerson, D. Franklin, S. and Wong S. (1992), "Parallel Genetic Search Via Clans," (with and Dinah Dickerson and Seok Wong), Intelligent Engineering Systems Through Artificial Neural Networks, vol 2 (Proceedings of the Artificial Neural Networks in Engineering (ANNIE '92) Conference), C.H. Dagli et al eds, New York: ASME Press, 927-935.
Maes, Pattie (1990). "How to do the right thing." Connection Science. (1):3.
Maes, Pattie (1991a), Designing Autonomous Agents, Cambridge MA: MIT Press.
Maes, Pattie (1994). "Agents that Reduce Work and Information Overload." Communications of the ACM, Vol. 37, No. 7. 31-40, 146, ACM Press, July.
Maes, Pattie (1991b), "A Bottom-up Mechanism for Behavior Selection in an Artificial Creature," From Animals to Animats, (editors J.-A. Meyer and S. W. Wilson), Cambridge MA: MIT Press, 238-246.
Neisser, Ulric (1993). "Without Perception, There Is No Knowledge: Implications for Artificial Intelligence." in R. G. Burton, ed., Natural and Artificial Minds, State University of New York Press.Pfeifer, R. (1996). Building "Fungus Eaters": Design principles of

autonomous agents. Proc. SAB-96, 3-12.

Song, H., Franklin, Stan and Negatu, A. (1996), "SUMPY: A Fuzzy Software Agent." in ed. F. C. Harris, Jr., Intelligent Systems: Proceedings of the ISCA 5th International Conference (Reno Nevada, June 1996) International Society for Computers and Their Applications - ISCA, 124-129.

Zhou, Lijia, and Franklin, Stan (1994). "Character Recognition Agents," Proceedings of Artificial Life IV, Cambridge, MA: MIT Press, 301-306.