Surfaces

PRACTICES FOR A MACHINE CULTURE: A CASE STUDY OF INTEGRATING CULTURAL THEORY AND ARTIFICIAL INTELLIGENCE

Phoebe Sengers

Volume 8, 1999

HUMANITÉS ET INFORMATIQUE : QUI A LES COMMANDES ?
HUMANITIES AND COMPUTING: WHO'S DRIVING?

URI: https://id.erudit.org/iderudit/1065079ar
DOI: https://doi.org/10.7202/1065079ar

See table of contents

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

1188-2492 (print)
1200-5320 (digital)

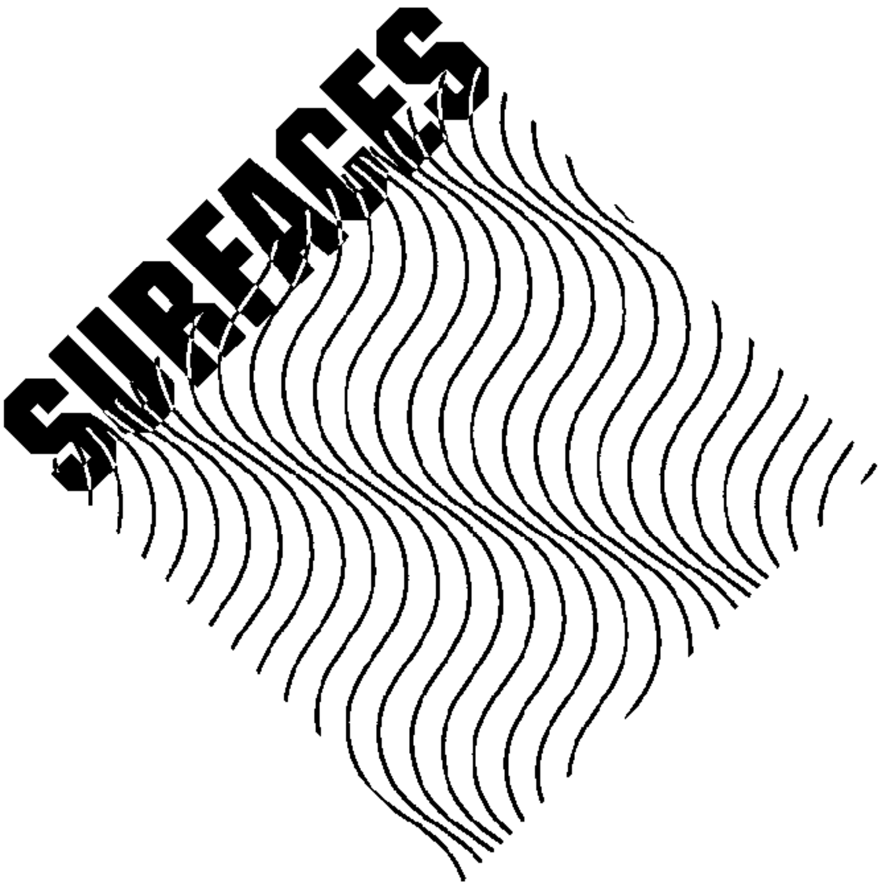
Explore this journal

Cite this article

Sengers, P. (1999). PRACTICES FOR A MACHINE CULTURE: A CASE STUDY OF INTEGRATING CULTURAL THEORY AND ARTIFICIAL INTELLIGENCE. *Surfaces*, *8*. https://doi.org/10.7202/1065079ar

Article abstract

In taking into account the contemporary development of technology, the author defends an integration of cultural theory and artificial intelligence research into what she calls cultural informatics. The author reviews the history of artificial intelligence from classical to what is now called alternative artificial intelligence research. A small but active community of researchers focusing on critical technical practices has developed in artificial intelligence research; among other items, they have questioned the traditional understanding of an agent as involved in purely logical and formal relations in order to take into account the embodiment of the agent.



PRACTICES FOR A MACHINE CULTURE

A CASE STUDY OF INTEGRATING CULTURAL THEORY AND ARTIFICIAL INTELLIGENCE

Phoebe Sengers

Remerciements:

SURFACES peut paraître grâce à un soutien des Presses de l'Université de Montréal, ainsi qu'à une aide matérielle importante (locaux, mobilier, services informatiques) de l'Université de Montréal.

Adresses:

SURFACES
Département de littérature comparée
Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal, Qc, Canada H3C 3J7

Tél.: (514) 343 5683
Fax: (514) 343 5684

Accès sur INTERNET
www.pum.umontreal.ca/revues/surfaces
Assistance technique et
questions éditoriales: **guedon@ere.umontreal.ca**
ISSN: 1188-2492

Soumission des textes

Vous pouvez soumettre vos articles, revues, notes, commentaires ou autres aux éditeurs par courrier électronique, par disquette ou par lettre. En ce qui concerne les notes bibliographiques, veuillez suivre le protocole du MLA. Nous vous demandons également de joindre à votre texte un court résumé de 3 à 4 lignes.

Tout article publié que l'auteur voudra modifier ultérieurement sera de nouveau évalué. Si accepté, il se substituera à la première version avec une nouvelle date et la mention v. 2.

Tout texte reste la propriété de son auteur. Néanmoins, **SURFACES** demande d'être citée à l'occasion de toute autre publication du texte en question.

Citation des textes

Pour toute citation, veuillez consulter le document source afin d'obtenir la dernière version du texte qui vous intéresse et vous baser sur la pagination flottante (/pp. xx/) plutôt que sur la foliation (pied de page). Si vous n'avez pas accès au courrier électronique, adressez-vous aux éditeurs qui pourront vous vendre une copie papier ou sur disquette.

RÉSUMÉ

Prenant en compte les développements technologiques récents, l'auteure défend une intégration des théories de la culture et de la recherche en intelligence artificielle dans ce qu'elle nomme l'informatique culturelle. L'auteure passe en revue l'histoire de la recherche en intelligence artificielle du moment classique aux recherches dites alternatives. Une communauté de chercheurs, encore petite mais dynamique, concentre ses efforts autour de pratiques techniques en intelligence artificielle qui soient critiques; parmi les enjeux que soulèvent ces chercheurs, l'auteure note la modification dans la compréhension de la notion d'agent, la faisant passer d'une notion impliquant des relations purement formelles et logiques à une notion impliquant la corporéité.

ABSTRACT

In taking into account the contemporary development of technology, the author defends an integration of cultural theory and artificial intelligence research into what she calls cultural informatics. The author reviews the history of artificial intelligence from classical to what is now called alternative artificial intelligence research. A small but active community of researchers focusing on critical technical practices has developed in artificial intelligence research; among other items, they have questioned the traditional understanding of an agent as involved in purely logical and formal relations in order to take into account the embodiment of the agent.

INTRODUCTION

We are early 21st-century humanity, the inheritors of industrialism, the progenitors of the information age. We live in a machine culture; in our daily lives, we are more and more surrounded by and interfaced with machines. We are no longer, like our ancestors, simply supplied by machines; we live in and through them. From our workplaces to our errands about town to our leisure time at home, human experience is to an unprecedented extent the experience of being interfaced with the machine, of imbibing its logic, of being surrounded by it and seeking it out: pager, cell phone, Palm Pilot, the latest version, the state of the art, the most advanced engineering.

Given that this is our cultural state, one of the most urgent questions we face as a society is the identification of practices which are adequate for intervening in its development. In this intimate machine-culture constellation, how can we decide what we ought to do? How should we as a society spend our resources? What interventions are possible, what are not possible, what are advisable, what had we better stay away from?

One candidate for a practice that could answer these questions is computer science, which has developed extensive practices for constructing computational machinery. Computer scientists understand well how machines can be built, what kinds of technology are possible, and what kinds could be possible if more effort were invested. They are trained to identify shortcomings in technology and to propose solutions to those shortcomings. In practice, they tend to have an intimate familiarity with the inner workings of machines of a sort which is difficult for non-technical-workers to develop. /p.5-6/

At the same time, computer science suffers a disciplinary amnesia to the machine's cultural context. Computer scientists are trained to focus on machinery, i.e. what can be done, but not on whether it should be done or how it will be applied. The computer remains a black box, within which computer scientists work and outside of whose impermeable boundary the rest of culture and society goes about its business. Questions of sociocultural implication are not answerable within this framework[1].

Another candidate for a practice to address machine culture is the cultural studies of science, which has developed extensive practices for analyzing technology in a cultural context. Cultural critics know how technology is taken up in and influences broader culture, as well as how cultural background — such as unconsciously held metaphors and philosophies — can encourage the development of certain forms of technology at the expense of others. Cultural critics also have access to tools for analyzing the political and material economies which enable particular forms of technologies and discourage others; they know the cultural pressure points.

At the same time, cultural studies is at a disadvantage in proposing new interventions in machine culture because, as Richard Doyle puts it, it has historically been a *consumer* of practices rather than a *producer* of ones. That is to say, cultural studies has the tendency to critique, rather than to generate new practices which /p. 6-7/

---

[1] The one major exception to this black-boxing is the field of human-computer interaction (HCI), which looks closely at the human context of computing.
See for example Brenda Laurel and S. Joy Mountford, *The Art of Human-Computer Interface Design*, Addison-Wesley, 1990. However, the more socioculturally interesting aspects of HCI generally remain ghettoized
there; HCI as a speciality serves as a reason for the nonspecialized to concentrate on other things.

respond to critique. As a result, it often lacks agency in the critiqued practices, being marginalized as a kind of disciplinary Cassandra. In addition, because cultural studies tends not to engage in the practices it criticizes, it frequently lacks the intimate (though not necessarily self-reflective) knowledge of those engaging in those practices, and may at times misunderstand them.

I believe the technical practices of computer science and engineering and the critical practices of cultural studies and the humanities both provide important ingredients to intervene in machine culture, but neither is sufficient alone. In order to be able to address contemporary human experience, we need science and the humanities to be combined into hybrid forms which can address the machinic and the human simultaneously. Squeezed in between the disciplines, we can already see these forms developing.

A CULTURAL INFORMATICS

At the confluence of computation and the humanities, there are already numerous hybrid practices. Computation itself is the object of humanities research: the history of computation, the sociology of computer use, cultural criticism of Artificial Life[2].

/p. 7-8/

[2] The history of computation: e.g. Paul Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America*, Cambridge, MA: MIT Press, 1997.
The sociology of computer use: e.g. J. Taylor and J. MacDonald, "The Effects of Electronic Interactions on Group and Individual Communication Processes," *Journal of Intelligent Systems*, vol 4, nos. 1-2, 1994, pp 113-132.
Cultural criticism of Artificial Life: e.g. Stefan Helmreich, *Silicon Second Nature: Culturing Artificial Life in a Digital World*, Berkeley, CA: University of California Press, 1998.

Computational tools are used for humanistic projects; humanists compose with word processors, send each other email, read the latest articles over the Web. At the same time, computational artefacts become essential research tools; automatic text analysis is used to support literary criticism, scholarly papers appear in hypertext, collaborative writing environments are used to co-author texts. And in conjunction with the adoption of computational tools, computational concepts are borrowed and adapted to humanist projects: chaos theory as a method of literary analysis, the cyborg as a model of subjectivity, the robot historian as first-person perspective[3].

These hybrid practices are an essential (and perhaps inevitable) response to machine culture. At the same time, the approaches outlined so far share one disadvantage: an underlying disciplinary split. Computation is seen from the outside, to be observed, analyzed, used, learned from. The development of computational tools, however, remains largely in the domain of computer scientists, to be informed by humanist wishes, to be intrigued by humanist appropriations, to be confused by humanist critique, but to be done using time-honored engineering methodologies.

But in a world where machinery is woven in to the fabric of our daily lives, it is, while useful, not enough to approach computation at an arm's length, to make it the object or pre-given tool /p. 8-9/

---

[3] Chaos theory as a method of literary analysis: see N. Katherine Hayles, *Chaos Bound: Orderly Disorder in Contemporary Literature and Science*, Ithaca: Cornell UP, 1990. The cyborg as a model of subjectivity: see Donna Haraway, "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century," in *Simians, Cyborgs and Women: The Reinvention of Nature*, New York; Routledge, 1991, pp.149-181. The robot historian as first-person perspective: see Manuel De Landa, *War in the Age of Intelligent Machines*, NY: Zone Books, 1991.

of the humanities. The humanities must not only observe, use, and critique computation, but also ingest it. Computing itself must become a humanist discipline.

What this does *not* mean is the simple use of humanist results in order to optimize computer programming, the development of analytic Shakespeare generators, the reduction of the humanities to what can be output by a computer program. Instead, humanist forms of computing can be a set of practices incorporating a critical, self-reflexive viewpoint into technical work, using the research strategies and values of the humanities, embodying those values and traditions in changing technologies that in turn change human lives. They are oriented towards and respect the full complexity of human experience in the world, rather than reducing that experience to simple rules in the traditions of the natural sciences. They carry a healthy scepticism about the origin and value of computational concepts and tools, but rather than reject them they reorient them. They realize that the term computer *science* is a historical term, originally used to establish the legitimacy of computing as a coherent and respectable discipline, now artificially limiting the full breadth of possible computational research. This is a *cultural informatics*. This is what a growing and hybrid group of artists, researchers, and critics already do.

Humanists have called for such successor science projects for years. The research tradition of a humanist computation, though somewhat buried under the overwhelming mass of traditional computer science, already exists and is gaining strength. It is generally unnoticed because humanistically-informed computing is still computing. It is specific, oriented towards a mostly scientific academic subculture, flying below the radar screens of the humanities.

/p. 9-10/

In this paper, I will also work specifically, looking at the confluence of cultural studies and Artificial Intelligence (AI). I will focus particularly on the subfield of *autonomous agents*, artificial creatures that 'live' in physical or virtual environments, capable of engaging in complex action without human control. While giving an overview of research in this field, I will explain how issues of subjectivity unconsciously arise, suggesting an entrypoint for cultural studies. I will lay out how cultural studies and agent research can be and are being synthesized, and look at the mostly unknown research tradition that already exists in this area. I will then connect the critical practices within AI to those in computer science and general, as well as complementary approaches to cultural informatics emerging from within the arts.

CASE STUDY: AI AND CULTURAL THEORY

Introduction to Autonomous Agents

One of the dreams of AI is the construction of autonomous agents, independent artificial beings. Rather than slavishly following our orders, or filling some tiny niche of activity that requires some aspect of intelligence (for example, playing chess), these artificial creatures would lead their own existences, have their own thoughts, hopes, and feelings, and generally be independent beings just as people or animals are. Autonomous agents would be more than useful machinery, they would be independent subjects.

This AI dream of mechanical creatures that are, in some sense, alive, can seem bizarre at first glance. It is therefore important to note that this is not an idea that is new in AI, but, as Simon Penny notes, the continuation of a tradition of anthropomor- /p. 10-11/

phization that extends back thousands of years[4]. In this sense, the AI dream is similar to the 'writing dream' of characters that ring true, to the 'painting dream' of images that seem to step out of the canvas, to the fantasies of children that their teddy bears are alive, and to many other Pygmalionesque dreams of human creations that begin to lead their own lives.

But there is certainly a sense in which AI brings a new twist to these old traditions. AI as a cultural drive needs to be seen in the context of post-industrial life, in which we are constantly surrounded by, interfaced with, and defined through machines. At its worst, AI adds a layer of seductive familiarity to that machinery, sucking us into a mythology of user-friendliness and humanity while the same drives of efficiency, predictability, quantifiability, and control lurk just beneath our perception.

But at its best, AI invokes a hope that is recognizable to humanists — that is invoked, in fact, by Donna Haraway in her « Cyborg Manifesto''[5]. This is the hope that, now that we are seemingly inescapably surrounded by technology, this technology can itself become hybridized and develop a human face. This version of the AI dream is not about the mechanistic and optimized reproduction of living creatures, but about the becoming-living of machines. The hope is that rather than forcing humans to interface with machines, those machines may learn to interface with us, to present themselves in such a way that they do not drain us of our humanity, but instead themselves become humanized.

/p. 11-12/

[4] Simon Penny. "Living Machines." *Scientific American*, September 1995. 150th Anniversary Issue.

[5] Donna Haraway, op. cit.

In the 1950's and early 1960's, this dream for AI, for good and for bad, was embodied in cybernetics. W. Grey Walter, for example, built small robots with rudimentary "agenty" behaviors[6]. He called his robots 'turtles;' they would roam around their environment, seeking light, finding food, and avoiding running into things. Later models could do some rudimentary associative learning.

But as cybernetics fell out of fashion, AI research began to focus more on the cognitive abilities an artificial agent might need to have higher-level intelligence, and less on building small, complete (if not so smart) robots. At least partially because the task of reproducing a complete creature has been so daunting, AI spent quite a few years focused on building individual intelligent capabilities, such as machine learning, speech recognition, story generation, and computer vision. The hope was that, once these capabilities were generated, they could be combined into a complete agent; the actual construction of these agents was often indefinitely deferred.

More recently, however, the field of autonomous agents has been enjoying a renaissance. The area of autonomous agents focuses on the development of programs that more closely approach representations of a complete person or creature. These agents are programs which engage in complex activity without the intervention of another program or person. Agents may be, for example, scientific simulations of living creatures, characters in an interactive story, robots who can independently explore their environment, or virtual 'tour guides' that accompany users on their travels on the World Wide Web[7]. From the early debacles /p. 12-

[6] W.Grey Walter. *The Living Brain*. W.W. Norton, 1963.

[7] For a scientific model, see e.g. Bruce Blumberg. "Action-Selection in Hamsterdam: Lessons from Ethology". In *Proceedings of the 3rd International Conference on the Simulation of Adaptive Behavior*. Brighton, 1994.

13/ of Microsoft Bob through the alternately loved and hated Microsoft Office paper clip to the commercial hits of Tamagotchi and Furby's, agents are making their way into the average Netizen's home and consciousness.

While these applications vary wildly, they share the idea that the program that underlies them is like a living creature in some important ways. Often these ways include being able to perceive and act on their (perhaps virtual) environment; being autonomous means they can make decisions about what to do based on what is happening around them and without necessarily consulting a human for help. Agents are also often imputed with rationality, which is defined as setting goals for themselves and achieving them reasonably consistently in a complex and perhaps hostile environment.

Agent as Metaphor

The definition of what exactly is and is not an agent has at times been the source of vehement controversy in the field. Mostly these controversies revolve around the fact that any strictly /p. 13-14/

For interactive characters, see e.g. Joseph Bates. "The Role of Emotion in Believable Agents". Technical Report CMU-CS-94-136, Carnegie Mellon University, 1994. Also appears in *Communications of the ACM*, Special Issue on Agents, July 1994.
For robots, see e.g. Reid Simmons, Richard Goodwin, Karen Zita Haigh, Sven Koenig, and Joseph O'Sullivan. "A Modular Architecture for Office Delivery Robots". In W. Lewis Johnson, editor, *Proceedings of the First International Conference on Autonomous Agents*, pages 245-252, NY, February 1997. ACM Press.
For virtual tour guides, see e.g. Thorsten Joachims, Dayne Freitag, and Tom Mitchell. "WebWatcher: A Tour Guide for the World Wide Web". *In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97)*, August 1997.

formal definition of agenthood tends to leave out such well-beloved agents as cats or insects, or include such items as toasters or thermometers that a lay person would be hard-pressed to call an agent. With some of the looser definitions of agents, for which the word 'agent' just seems to be a trendy word for 'program,' skeptics can be forgiven for wondering why we are using this term at all.

Here, I will take agenthood broadly to be a sometimes-useful way to frame inquiry into the technology we create. Specifically, agenthood is a metaphor we apply to computational entities we build when we wish to think of them in ways similar to the ways we understand living creatures. Calling a program an agent means the program's designer or the people who use it find it helpful or important or attractive to funders to think of the program as an independent and semi-intelligent coherent being. For example, when we think of our programs as agents we focus our design attention on 'agenty' attributes we would like the program to have : the program may be self-contained ; it may be situated in a specific, local environment ; it may engage in 'social' interactions with other programs or people[8]. When a program is presented to its user as an agent, we are encouraging the user to think of it not as a complex human-created mechanism but as a user-friendly, intelligent creature. If 'actually' some kind of tool, the creature is portrayed as fulfilling its tool-y functions by being willing to do the user's bidding[9]. Using the metaphor 'agent' for these applications lets us apply ideas about what living agents such as dogs, beetles, or bus drivers are like to the design and use of artificially-created programs.

/p. 14-15/

---

[8] I am indebted to Filippo Menczer for this observation.

[9] See: Jaron Lanier. "My Problem With Agents". In *Wired*, 4(11), November 1996.
J. MacGregor Wise. "Intelligent Agency." In *Cultural Studies*, 12 (3), 1998.

Agenthood in Classical and Alternative AI

But not all AI researchers agree on which conceptions of living agents are appropriate or useful for artificial agents. The past 15 years in particular have seen an at times spectacular debate between different strains of thought about the proper model of agent to use for AI research[10]. Rodney Brooks, for example, distinguishes between 'symbolically-grounded' and 'physically-grounded' agents[11]. These symbolically-grounded agents spent most of their time in abstract cogitation; their programs manipulate representations of the "real world" (for example, in database form), but rarely come into contact with that real world. Physically-grounded agents, on the other hand, manipulate and react to the environment itself without having external objects explicitly represented in their program code.

Philip Agre and David Chapman distinguish agents using 'plans-as-programs' from agents using 'plans-as-communication.' This is a distinction based on the relative importance of internally-determined planned-out activity versus a more improvised, moment-by-moment immersion in environmental circumstances. Agents that use plans as programs are heavily invested in their internal representation of action; they engage in abstract, hierarchical planning of activity before engaging in it (often including formal proofs that the plan will fulfill the goal the agent is given). Agents that use plans as communication see plans as a convenience but not a necessity. They are designed to take advantage of an action loop with respect to their /p. 15-16/

---

[10] See e.g. *Cognitive Science*, January - March 1993. Volume 17. No. 1. Special Issue on Situated Action

[11] Rodney A. Brooks. "Elephants Don't Play Chess". In Pattie Maes, editor, *Designing Autonomous Agents*. MIT Press, Cambridge, MA, 1990.

environment and may only refer to plans as ways to structure common activities[12].

Another common distinction is between situated and cognitive agents. Situated agents are thought of as embedded within an environment, and hence highly influenced by their situation and physical make-up. Cognitive agents, on the other hand, engage in most of their activity at an abstract level and without reference to their concrete situation.

Each of these distinctions is not independent of the others. When looking at such classification attempts at a whole, a distinct theme emerges. AI research in general can be understood as involving two major trends in thinking: a main stream often termed *classical* AI (also known as Good Old-Fashioned AI, cognitivistic AI, symbolic cognition, top-down AI, knowledge-based AI, etc.) and an oppositional stream we can term *alternative* AI (also known as new AI, nouvelle AI, ALife, behavior-based AI, reactive planning, situated action, bottom-up AI, etc.)[13]. Not every AI system neatly falls into one or the other category — in fact, few can be said to be pure, unadulterated representatives of one or other. But each stream represents a general trend of thinking about agents that a significant number of systems share. /p. 16-17/

---

[12] Philip E. Agre and David Chapman. "What Are Plans For?" In Pattie Maes, editor, *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, pages 17-34. MIT Press, Cambridge, MA, 1990.

[13] For similar analyses, see e.g. the following: Luc Steels. "The Artificial Life Roots of Artificial Intelligence". In *Artificial Life*, 1(1-2):75--110, 1994.
Francisco J. Varela, Evan Thompson, and Eleanor Rosch. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, Cambridge, MA, 1991.
Brooks, op. cit.
Donald A. Norman. "Cognition in the Head and in the World: An Introduction to the Special Issue on Situated Action." *Cognitive Science*, 17:1-6, 1993.

For AI researchers, the term classical AI refers to a class of representational, disembodied, cognitive agents, based on a model that proposes, for example, that agents are or should be fully rational and that physical bodies are not fundamentally pertinent to intelligence. The more extreme instances of this type of agent had their heyday in the 60's and 70's, under a heady aura of enthusiasm that the paradigms of logic and problem-solving might quickly lead to true AI. One of the earliest examples of this branch of AI is Allen Newell and Herbert Simon's GPS, the somewhat optimistically titled "general problem solver." This program proceeds logically and systematically from the statement of a mathematical-style puzzle to its solution[14]. Arthur Samuel's checker player, one of the first programs that learns, attempts to imitate intelligent game-playing by learning a polynomial function to map aspects of the current board state to the best possible next move[15]. Terry Winograd's SHRDLU maintains a simple representation of blocks lying on a table, and uses a relatively straightforward algorithm to accept simple natural language commands to move the virtual blocks[16]. While the creators of these programs often had more subtle understandings of the nature of intelligence, the programs themselves reflect a hope that simple, logical rules might underlie all intelligent behavior, and that if we could discover those rules we might soon achieve the goal of having intelligent machinery.

/p.17-18/

---

[14] Allen Newell and Herbert A. Simon. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ, 1972.

[15] Arthur L. Samuel. "Some Studies in Machine Learning Using the Game of Checkers." In Edward A. Feigenbaum, editor, *Computer and Thought*. AAAI Press, Menlo Park, 1995.

[16] Terry Winograd. *Understanding Natural Language*. Academic Press, New York, 1972.

But the classical model, while allowing programs to succeed in many artificial domains which humans find difficult, such as chess, unexpectedly failed to produce many behaviors humans find easy, such as vision, navigation, and routine behavior. The recognition of these failures has led to a number of responses in the 80's and 90's. Some researchers — most notably Winograd, who wrote an influential book with Fernando Flores on the subject[17] — have decided that the intellectual heritage of AI is so bankrupt they have no choice but to leave the field. By far the majority of AI researchers have remained in a tradition that continues to inherit its major research framework from classical AI, while expanding its focus to try to incorporate traditionally neglected problems (we might call this 'neo-classical AI'). A smaller but noisy group has split from classical AI, claiming that the idea of agents that classical AI tries to promote is fundamentally wrong-headed.

These researchers, who we will here call alternative AI, generally believe that the vision of disembodied, problem-solving minds that explicitly or implicitly underlies classical AI research is misguided. Alternative AI focuses instead on a vision of agents as most fundamentally nonrepresentational, reactive, and situated. Alternative AI, as a rubric, states that agents are situated within an environment, that their self-knowledge is severely limited, and that their bodies are an important part of their cognition.

Agent Technology as Theory of Subjectivity

The dialogue and debate between these two types of agents is not only about a methodology of agent-building. An underlying source of conflict is about which aspects of being human /p. 18-19/

[17] Terry Winograd and Carlos F. Flores. *Understanding Computers and Cognition: A New Foundation for Design.* Ablex Pub. Corp., Norwood, NJ, 1986

are most essential to reproduce. Classicists do not deny that humans are embodied, but the classical technological tradition tends to work on the presupposition that problem-solving rationality is one of the most fundamental defining characteristic of intelligence, and that other aspects of intelligence are subsidiary to this one. Likewise, alternativists do not deny that humans can solve problems and think logically, but the technology they build is based on the assumption that intelligence is inherent in the body of an agent and its interactions with the world; in this view, human life includes problem-solving, but is not a problem to be solved.

It is in these aspects of AI technology — ones that are influenced by and in turn influence the more philosophical perspectives of AI researchers — that we can uncover, not just the technology of agents, but also theories of agenthood. Two levels of thought are intertwined in both these approaches to AI: (1) the level of day-to-day technical experience, what works and what doesn't work, which architectures can be built and which can't; and (2) the level of background philosophy — both held from the start and slowly and mostly unconsciously imbibed within the developing technical traditions — which underlies the way in which the whole complex and undefined conundrum of recreating life in the computer is understood. Running through and along with the technical arguments are more philosophical arguments about what human life is or should be like, how we can come to understand it, what it means to be meaningfully alive.

The argument is straightforward: if agents are metaphors that are used to design programs that are in some sense like people, then the way we build agents will depend on and in turn reveal a great deal about what we think people are like. This means AI includes not only conflicting theories of technology but also, implicitly, conflicting theories of subjectivity. Classical AI technology /p. 19-

20/ is based on a model of subjectivity as essentially representational, rational, and disembodied. Alternative AI technology presupposes that it is essentially reactive, situated, and embodied.

These two categories can be clearly seen within AI research. Within that research community, they are generally seen as arising from certain tensions in technical practice itself. But these categories should be familiar to cultural theorists from a quite different context; they directly correspond to *rational* (or Enlightenment) and *schizophrenic* (or postmodern) subjectivity[18].

Rational subjectivity is based on the Cartesian focus on logical thought: the mind is seen as separated from the body, it is or should be fundamentally rational, and cognition divorced from emotion is the important part of experience. This model has overarching similarities with, for instance, Allen Newell's theory of Soar, which describes an architecture for agents that grow in knowledge through inner rational argumentation[19]. Most models built under Soar are focused on how this argumentation should take place, leaving out issues of perception and emotion (though there are certainly exceptions[20]. /p. 20-21/

---

[18] This idea is a more common observation among cultural theorists who study AI. See, for example: Will Barton. "Letting Your Self Go: Hybrid Intelligence, Shared Cognitive Space and Posthuman Desire." Presented at Virtual Futures, Coventry, UK., 1995.
Jos de Mul. "Networked Identities: Human Identity in the Digital Era". In Michael B. Roetto, editor, *Proceedings of the Seventh International Symposium on Electronic Art*, pages 11-16, Rotterdam, 1997.

[19] Allen Newell.*Unified Theories of Cognition*. Harvard University Press, Cambridge, Massachusetts, 1990.

[20] See e.g. Douglas J. Pearson, Scott B. Huffman, Mark B. Willis, John E. Laird, and Randolph M. Jones. "Intelligent Multi-level Control in a Highly Reactive

The development of the notion of schizophrenic subjectivity is based on perceived inadequacies in the rational model, and is influenced by but by no means identical to the psychiatric notion of schizophrenia. While rational subjectivity presupposes that people are fundamentally or optimally independent rational agents with only tenuous links to their physicality, schizophrenic subjectivity sees people as fundamentally social, emotional, and bodily. It considers people to be immersed in and to some extent defined by their situation, the mind and the body to be inescapably interlinked, and the experience of being a person to consist of a number of conflicting drives that work with and against each other to generate behavior. In AI, this form of subjectivity is reflected in Brooks's subsumption architecture, in which an agent's behavior emerges from the conflicting demands of a number of loosely coupled internal systems, each of which attempts to control certain aspects of the agent's body based almost entirely on external perception rather than on internal cogitation[21].

Each class of agent architectures closely parallels a kind of subjectivity. Just as alternative AI has arisen in an attempt to address flaws in classical AI, the concept of schizophrenic subjectivity has arisen in response to perceived flaws in the rational model's ability to address the structure of contemporary experience. Each style of agent architecture shows a striking similarity to a historical model of subjectivity that cultural theorists have identified.

/p. 21-22/

Domain." In *Proceedings of the International Conference on Intelligent Autonomous Systems*, Pittsburgh, PA, 1993.

[21] Rodney Brooks. "A Robust Layered Control System for a Mobile Robot". *IEEE Journal of Robotics and Automation,* RA-2:14-23, April 1986.

This close relationship between a technical debate in a subfield of computer science and philosophical trends in Western culture as a whole generally comes as a surprise to technical workers. But the connection is obvious to cultural theorists. AI researchers are also human beings, and as such inhabit and are informed by the broader society that cultural theorists study. From this point of view, AI is simply one manifestation of culture as a whole. Its technical problems are one specific arena where the implications of ideas that are rooted in background culture are worked out.

But if AI is fundamentally embedded in and working through culture, then cultural studies and AI may have a lot to say to each other. Specifically, cultural theorists have spent a lot of time thinking about and debating subjectivity. AI researchers have spent a lot of time thinking about and debating architectures for autonomous agents. Once these two are linked, each body of work can be used to inform the other. If agents use a particular theory of subjectivity, then we can use ideas about this theory to inform our work on agents. And if agents are a manifestation of a theory of subjectivity, then studying these agents can give us a better idea of what that theory means. This raises the possibility that cultural studies and AI can form a strategic alliance.

CULTURAL STUDIES AND AI IN THE AGE OF THE SCIENCE WARS

Certainly cultural studies has not turned a blind eye to the ascendancy of science and technology in contemporary culture. The last 15 or 20 years has seen an explosion of research analyzing the complex relationships between science and the rest of culture. This, at least theoretically, lays the groundwork for a potential collaboration between science studies and science.

/p. 22-23/

Science studies, after all, examines culturally-based metaphors that inform scientific work, and thereby often uncovers deeply-held but unstated assumptions that underlie it. Scientists are also generally interested in understanding the forces, both conscious and unconscious, that can shape their results. If there are ways in which they can better understand the phenomena they study or build the technology they want to create, they are all ears. In this respect, as Evelyn Fox Keller points out, the insights of science studies can contribute great value to science's self-understanding[22].

At the same time, many practitioners of science studies are deeply interested in science as it is actually practiced on a day-to-day level. This means scientists, with their in-depth personal experience of what it means to do scientific work, are privy to perspectives that can enrich the work of their science studies counterparts. Science studies simply is not possible without science, and an important component of it is an accurate reflection of the experiences of scientists themselves.

A Siege Mentality

With all the advantages that cooperation could bring, you might think that science and science studies would be enthusiastic partners on the road to a shared intellectual enterprise. Alas, the practitioners of science studies and many of their hapless subjects know that that is far from the case. Productive exchanges between cultural critics and scientists interested in the roots of their work are hampered by the disciplinary divide between them. This divide blocks cultural critics from access to a complete understanding of the process and experience of doing science, which can /p. 23-24/

[22] Evelyn Fox Keller. *Reflections on Gender and Science*. Yale University Press, New Haven, 1985.

degrade the quality of their analyses and may lead them to misinterpret scientific practices. At the same time, scientists have difficulty understanding the context and mindset of critiques of their work, making them unlikely to consider such critiques seriously or realize their value for their work, potentially even leading them to dismiss all humanistic critiques of science as fundamentally misguided[23].

This feedback loop of mutual misunderstanding has grown into a new tradition of mutual kvetching. Cultural critics may complain that scientists unconsciously reproduce their own values in their work and then proclaim them as eternal truth. They may feel that scientists are not open to criticism because they want to protect their high (relative to the humanities') status in society. Simultaneously, scientists sometimes complain that cultural critics are absolute nihilists who do not believe in reality and equate science with superstition[24]. They fear that cultural critics undermine any right that science has as a source of knowledge production to higher status than, say, advertising. Finally, both sides complain incessantly — and correctly — of being cited, and then judged, out of context.

The unfortunate result of this situation is a growing polarization of the two sides. In the Science Wars, pockets of fascinating interdisciplinary exchanges and intellectually illuminating debate are sadly overwhelmed by an overall lack of mutual understanding and accompanying decline of goodwill. While most par- /p.24-25/

[23] For a case in point, see Paul R. Gross and Norman Levitt. *Higher Superstitions: The Academic Left and Its Quarrels with Science*. Johns Hopkins University Press, Baltimore, 1994.

[24] This is exacerbated by the fact that the notion of 'reality' used by many scientists in their criticism of science studies does not bear much relation to the long and deep tradition of the usage of that term in cultural studies of science.

ticipants on both sides of the divide are fundamentally reasonable, communication between them is impaired when both sides feel misunderstood and under attack. This siege mentality not only undermines the possibility for productive cooperation ; with unfortunate frequency, it goes as far as cross-fired accusations of intellectual bankruptcy in academic and popular press and nasty political battles over tenure. These unpleasant incidents not only help no one but also obscure the fact that *both* the academic sciences *and* the humanities are facing crises of funding in an economy that values quick profit and immediate reward over a long-term investment in knowledge. In the end, neither science nor science studies benefits from a situation best summed up from *both* sides by Alan Sokal's complaint : « The targets of my critique have by now become a self-perpetuating academic subculture that typically ignores (or disdains) reasoned criticism from the outside''[25].

Science Wars, AI Skirmishes

While most scientists remain blissfully unaware of the Science Wars, they are not unaffected by them. Within AI, the tension between the self-proclaimed defenders of scientific greatness and the self-identified opponents of scientific chauvinism is worked out under the table. In particular, the sometimes tendentious clashes between classical and alternative AI often reflect arguments about science and the role of culture in it.

/p. 25-26/

---

25 Sokal happens to be a physicist complaining about science studies, but this quote works just as aptly to summarize the complaints made the other way around. Alan Sokal "A Physicist Experiments with Cultural Studies" In *Lingua Franca*, pages 62-64, May-June 1996.

This can be seen most clearly in a rather unusual opinion piece that appeared several years ago in the *AI Magazine*[26]. The remarkable rhetoric of this essay in a journal more often devoted to the intricacies of extracting commercially relevant information from databases may be appreciated in this excerpt:

> Once upon a time there were two happy and healthy babies. We will call them Representation Baby (closely related to Mind Baby and Person Baby) and Science Baby (closely related to Reality Baby).
>
> These babies were so charming and inspirational that for a long time their nannies cared for them very well indeed. During this period it was generally the case that ignorance was pushed back and human dignity increased. Nannies used honest, traditional methods of baby care which had evolved during the years. Like many wise old folk, they were not always able to articulate good justifications for their methods, but they worked, and the healthy, happy babies were well growing and having lots of fun.
>
> Unfortunately, some newer nannies haven't been so careful, and /p. 26-27/

[26] Patrick J. Hayes, Kenneth M. Ford, and Neil Agnew. "On Babies and Bathwater: A Cautionary Tale". In AI Magazine, 15(4):15-26, 1994.

> the babies are in danger from their zealous ways. We will focus on two nannies who seem to be close friends and often can be seen together - Situated Nanny (called SitNanny for short) and Radical Social Constructivist Nanny (known to her friends as RadNanny) (15)[27].

A little decoding is in order for those not intimately aware of both the AI debates and the Science Wars. "SitNanny" represents situated action, a brand of alternative AI that focuses its attention on the way in which agents are intimately related to, and cannot be understood without, their environment. "RadNanny," as is immediately clear to even the most naive science studies aficionado, is the embodiment of the cultural studies of science, social constructivism being the belief that science, like every other human endeavor, is at least partially a product of sociocultural forces (the 'radical' here functions as little more than an insult, but implies that science is *purely* social, i.e. has absolutely no relationship to any outside reality).

Having broken the code, the implication of this excerpt is clear: everything in AI was going fine as long as we thought about things in terms of science and knowledge representation, one of the core terms of classical AI. Of course, this science was not always well-thought-out, but it was fundamentally good. That is, until that dastardly alternative AI came along with cultural studies in its tow and threatened nothing less than to *kill the babies*. /p. 27-28/

[27] This excerpt cannot, however, carry the full force of the original, which contains several full-page 19th-century woodcuts displaying suffering babies and incompetent or evil nannies (labeled, for example, "The Notorious RadNanny Looking For Babies").

Now any cultural critic worth his or her salt will have some choice commentary on a story in which the positive figures are all male babies living the life of leisure, and the negative figures all lower-class working women[28]. But the really interesting rhetorical move in this essay is in the alignment of the classical-alternative AI debate with the Science Wars. Classical AI, we learn, is good science. Alternative AI, while having some good ideas, is dangerous, among other reasons because it is watering down science with other ideas: "concepts from fringe neurology, sociology, ethnomethodology, and political theory; precomputational psychological theory; and God knows what else" (19). Alternative AI is particularly dangerous because it believes that agents cannot be understood without reference to their environment. Hence, it is allied with the "cult" (20) of science studies, which believes that scientists cannot be understood without reference to their sociocultural environment.

Since the majority of their audience presumably has little awareness of science studies, the authors are happy to do their part for interdisciplinary awareness by explaining what it is. They state, in a particularly nice allusion to 1950's anti-Communist hysteria, that science studies aims at nothing less than to "reject the entire fabric of Western science" (15). Science studies, we are informed, believes "that all science is arbitrary and that reality is merely a construction of a social game" (23). In the delightful tradition of the Science Wars, several quotations are taken out of context to prove that cultural critics of science believe that science is merely an expendable myth.

/p. 28-29/

---

[28] One must presume that the authors were aware of this and did their best to raise cultural critics' hackles.

The statements Hayes et al. make are simply inaccurate descriptions of science studies. In reality, science studies tends to be agnostic on such questions as the arbitrariness of science and on the nature of reality, to which science studies generally does not claim to have any more access than science does. When science studies *does* look into these issues it does so in a much more subtle and complex way than simply rejecting or accepting them.

But what is more important than these factual inaccuracies is that the article promotes the worst aspects of the Science Wars, since the very *tone* of the article is chosen to preclude the possibility of productive discussion. Science studies is simply dismissed as ludicrous. If uninformed scientists reading the article have not by the end concluded that science studies is an evil force allied against them, with alternative AI its unfortunate dupe, it is certainly not for lack of trying

AI IN CULTURE, AI AS CULTURE

But is it really true that science studies is an enemy of AI? After all, no one disputes that AI is, among other things, a social endeavor. Its researchers are undeniably human beings who are deeply embedded in and influenced by the social traditions in which they consciously or unconsciously take part, including but by no means limited to the social traditions of AI itself. It seems that taking these facts seriously might not necessarily damage AI, but could even help AI researchers do their work better.

In this section, we will buck the trend of mutual disciplinary antagonism by exploring the potential of what /p. 29-30/

former agent researcher Philip Agre calls *critical technical practices*[29]. A critical technical practice is a way of actually doing AI which incorporates a level of reflexive awareness of the kind espoused by science studies. This may include awareness of the technical work's sociocultural context, its unconscious philosophies, or the metaphors it uses. We will look at various AI researchers who have found ideas from cultural studies helpful in their technical work.

A Short History of Critical Technical Practices

From the rather heated rhetoric of the Science Wars, one might be tempted to think that science and science studies have nothing of value to share with each other. Often, voices on the 'pro-science' side of the debate say that the cultural studies of science has no right to speak about science because only scientists have the background and ability to understand what science is about and judge it appropriately. At the same time, the 'pro-culture' side of the debate may feel that scientists neither know about nor care to ameliorate the social effects of their work.

These attitudes can only be maintained by studiously avoiding noticing the people who are *both* scientists *and* cultural critics. Gross and Levitt's influential onslaught against science studies[30], for example, argues that cultural critics are irresponsible and dangerous because they are ignorant of the science they criticize. This argument is made easier by counting interdisciplinarians who do *both* science *and* cultural studies as (good, responsible) scientists and not as (bad, irresponsible) cultural /p. 30-31/

29 Philip E. Agre. *Computation and Human Experience*. Cambridge University Press, Cambridge, UK, 1997.

30 Gross and Levitt, op. cit.

critics (the question of why those scientists would find it interesting or even fruitful to keep such unseemly company is left unanswered). And in an exhaustive survey of every important figure in cultural studies, some of the most influential 'culturalist scientists' are left out altogether. A glaring omission is Richard Lewontin, whose influential books on the cultural aspects of biology are the sidelight to an illustrious career as a geneticist[31].

Similarly, the hypothesis that scientists do not know or care about the effects of their work is contradicted by the work of Martha Crouch[32]. Crouch is a botanist who, after many years of research, noticed that the funding of botany combined in practice with the naive faith of scientists in their own field to completely undermine the idealistic goals of plant scientists themselves. Crouch determined to help scientists such as herself achieve their own stated goals of, for example, feeding the hungry, by adding to their self-understanding through the integration of cultural studies with botany.

But, to be fair, much of the work integrating science with science studies may be invisible to both cultural critics themselves and the scientists whose form of intellectual output seems to largely be attacks on those on the other side of the great intellectual /p. 31-32/

---

[31] See, for example, Richard Levins and Richard C. Lewontin. *The Dialectical Biologist*. Harvard University Press, Cambridge, MA, 1985.
Richard C. Lewontin, Steven Rose, and Leon J. Kamin. *Not In Our Genes: Biology, Ideology, and Human Nature*. Pantheon Books, New York, 1984.
Richard C. Lewontin. *Biology as Ideology: The Doctrine of DNA*. Harper Perennial, New York, 1991.
For Lewontin's roasting response to Gross and Levitt, see Richard C. Lewontin. À la Recherche du Temps Perdu." In *Configurations*, 3(2):257-265, 1995.

[32] Martha L. Crouch. "Debating the Responsibilities of Plant Scientists in the Decade of the Environment. " In *The Plant Cell*, pages 275--277, April 1990.

divide. This is because scientists who are actually using culturalist perspectives in their work generally address that work to their scientific subcommunity, rather than to all of science and science studies as a whole. And in work that is addressed to a technical subfield, it is usually not particularly advantageous to mention that one's ideas stem from the humanities, particularly if they come from such unseemly company as hermeneutics, feminism or Marxism.

Here, we will uncover the history of the use of culturalist perspectives within AI as a part of technical work. It turns out that within AI, the use of cultural studies perspectives is not just a couple of freak accidents traceable to a few lone geniuses and / or lunatics. Rather, there is a healthy if somewhat hidden tradition of a number of generations of AI researchers who have drawn inspiration from the humanities in ways that have had substantial impact on the field as a whole. We look at both how cultural studies was found to be useful, and the concrete methods various researchers have used to combine the fields.

Winograd and Flores

Terry Winograd is one of the first and certainly one of the most notorious in his usage of critical theory to analyze AI from the AI researcher's point of view. As mentioned in the review of classical AI, Winograd was a well-known researcher into the machine generation of human language. In collaboration with economist Fernando Flores, Winograd started exploring the implications of Heideggerian philosophy for AI. Unexpectedly, this resulted in Winograd's wholesale rejection of AI as intellectually bankrupt.

/p. 32-33/

In *Understanding Computers and Cognition*, Winograd and Flores analyze AI as a continuation of the analytic tradition[33]. AI's investment in this tradition, they conclude, is so great that it cannot address what they consider to be fundamental attributes of intelligence. Their critique is based on the Heideggerian notion that people approach the world from a set of prejudices that cannot be finitely articulated. If these prejudices cannot be finitely articulated, then they cannot be explicitly represented in machinery; any machinic representation of subjectivity will therefore necessarily leave out some of the complex background knowledge with which people approach real-world situations. This means that AI is able to solve limited, formal problems, but cannot attain true intelligence because "[t]he essence of intelligence is to act appropriately when there is no simple pre-definition of the problem or the space of states in which to search for a solution" (98). Winograd and Flores argue that instead of making computers that can communicate with us, we should make computers a means to aid communication between people.

While Winograd and Flores's arguments certainly made a splash in the field, it must be honestly stated that they probably did not cause too many scientists to leave AI (and they were not intended to). The basic flaw from this perspective in the argument is that it forces AI researchers to choose between believing in Heidegger and believing in AI. One can hardly blame them if they stay with the known evil.

What is interesting to those who remain in AI, however, is Winograd and Flores's methodology for combining a critical perspective with AI. Winograd and Flores analyze the limitations of AI that stem from its day-to-day methodologies. /p. 33-34/

---

[33] Winograd and Flores, op. cit.

When they find those constraints to exclude the possibility of truly intelligent behavior, they decide instead to start building systems in which those constraints become strengths. In other words, they decide that artificial systems necessarily have certain characteristics of rigidity and literalness, then ask themselves what sorts of social situations could be aided by a rigid, literal system. They then build a system that is an enforcer of social contracts in certain, limited situations where they feel it is important that social agreements be clearly delineated and agreed upon. Specifically, the system articulates social agreements within work settings, so that workers are aware of who has agreed to do what. This new system is designed to be useful precisely because of the things that were previously limitations. Winograd and Flores, then, use cultural studies to inform technical development by finding constraints in its methodologies, and then using those constraints so that they become strengths.

Suchman

Lucy Suchman is an anthropologist who, for a time, studied AI researchers and, in particular, the ideas of 'planning'[34]. Planning is an area of AI that is, at its most broad, devoted to deciding what to do. Since this broad conception does not really help you sink your teeth into the problem, a more limited notion has been generally used in AI. This concept of planning is a type of problem-solving where an agent is given a goal to achieve in the world, and tries to imagine a set of actions that can achieve that goal, generally by using formal logic.

/p. 34-35/

34 Lucy A. Suchman. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge, 1987.

Suchman noticed that the ideas of planning were heavily based on largely Western notions of, among other things, route planning. She then asked herself what kind of 'planning' you would have if you used the notions of a different society. By incorporating perspectives from Micronesian society, she came up with the concept of 'situated action,' which you may remember as the butt of ridicule in Hayes et al.'s "On Babies and Bathwater."

Situated action's basic premise is to generate behavior on the fly according to the local situation, instead of planning far ahead of time. Although Suchman herself made no claims to technical fame, her ideas became influential among AI researchers who were working on similarly-motivated technology (see below), becoming an important component in an entire subfield AI researchers now either love or hate, but generally cannot ignore. Her methodology, in sum, is to notice the culture-boundedness of a particular metaphor ("planning") that informs technical research, then ask what perspectives a very different metaphor might bring to the field instead. The point in her work is not that Western metaphors are 'wrong' and non-Western ones are 'right,' but that new metaphors can spawn new machinery that might be interesting in different ways from the old machinery.

Chapman

David Chapman was a graduate student at MIT when together with Agre, whose work is described separately below, he developed an agent architecture that was heavily influenced by Suchman's ideas, as well as by ethnomethodology[35]. Chapman's contribution in this history of interdisciplinary methodologies in AI is his /p. 35-36/

35 David Chapman. *Vision, Instruction, and Action*. PhD thesis, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 1990.

articulation of the value of 'ideas' — as opposed to proofs or technical implementation — in technical practice.

Chapman argues that some of the most interesting papers in AI do not make technical contributions in any strict sense of the term — i.e., that the best methodology for AI is not necessarily that of empirical natural science. "[Some of the best] papers prove no theorems, report no experiments, offer no testable scientific theories, propose technologies only in the most abstract terms, and make no arguments that would satisfy a serious philosopher. [...] [Instead, these] papers have been influential because they show us powerful ways of thinking about the central issues in AI" (214). Suchman's anthropological work in AI is a living example in Chapman's work of such an influential idea.

Agre

Of all AI researchers, Agre has probably done the most extensive and explicit integration of critical viewpoints with AI technology. In his thesis, for example, Agre integrates ethnomethodology with more straightforward AI techniques[36]. He uses ideas from ethnomethodology both to suggest what problems are interesting to work on (routine behavior, instead of expert problem-solving) and to suggest technical solutions (deictic, or subjective representation instead of objective representation).

Together, Chapman and Agre develop novel techniques for building agents which are based on a new conceptualization of what it means to be an agent. This conceptualization has roots in Winograd's Heideggerian analysis of AI, and is also /p. 36-37/

36 Philip E. Agre. *The Dynamic Structure of Everyday Life*. PhD thesis, Massachusetts Institute of Technology Artificial Intelligence Laboratory, Cambridge, MA, 1988.

deeply influenced by ethnomethodology, particularly Garfinkel and Suchman's work described above. Chapman and Agre reject the idea that problem-solving is central to agenthood, and instead see agenthood as process, engaging in a rich set of interactions with other agents and the physical world.

> The world of everyday life... is not a problem or a series of problems. Acting in the world is an ongoing process conducted in an evolving web of opportunities to engage in various activities and contingencies that arise in the course of doing so... The futility of trying to control the world is, we think, reflected in the growing complexity of plan executives. Perhaps it is better to view an agent as *participating* in the flow of events. An embodied agent must *lead a life*, not *solve problems*[37].

This re-understanding of the notion of agent has been an important intellectual strand in alternative AI's reconceptualization of agent subjectivity.

In recent work, Agre has distilled his approach to combining philosophy, critical perspectives, and concrete technical work into an articulated methodology for critical technical practices per se. Agre sees critical reflection as an indispensable tool in technical work itself, because it helps technical researchers to understand in a deep sense what technical impasses are trying to tell /p. 37-38/

[37] Agre and Chapman, op. cit., 20.

them. He sums up his humanistic approach to AI with these postulates:

> 1. AI ideas have their genealogical roots in philosophical ideas. 2. AI research programs attempt to work out and develop the philosophical systems they inherit. 3. AI research regularly encounters difficulties and impasses that derive from internal tensions in the underlying philosophical systems. 4. These difficulties and impasses should be embraced as particularly informative clues about the nature and consequences of the philosophical tensions that generate them. 5. Analysis of these clues must proceed outside the bounds of strictly technical research, but they can result in both new technical agendas and in revised understandings of technical research itself[38].

Humanists will recognize Agre's methodology as hermeneutics; it is a kind of interpretation that goes beyond surface appearances to discover deeper meanings. For Agre, purely technical research is the surface manifestation of deeper philosophical systems. While it is certainly possible for technical traditions to proceed without being aware of their philosophical bases, technical impasses provide clues that, when properly interpreted, can reveal /p. 38-39/

38 Philip E. Agre. "The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project". In *Stanford Humanities Review*, 4(2), 1995. Special issue — Constructions of the Mind: Artificial Intelligence and the Humanities.

the philosophical tensions that lead to them. If these philosophical difficulties are ignored, chances are that technical impasses will proliferate and remain unresolved. If, however, they are acknowledged, they can become the basis for a new and richer technical understanding.

In *Computation and Human Experience*, Agre develops a methodology for integrating AI and the critical tradition through the use of deconstruction[39]. This works as follows:

> 1. Find a metaphor that underlies a particular technical subfield. An example of such a metaphor is the notion of disembodiment that underlies classical AI.
> 2. Think of a metaphor that is the *opposite* of this metaphor. The opposite of disembodied agents would be agents that are fundamentally embodied.
> 3. Build technology that is based on this opposite metaphor. Embodied agents are an essential component of Rod Brooks's ground-breaking work at the core of alternative AI, as described above[40]. This technology will inevitably have both new constraints and new possibilities when compared to the old technology. /p. 39-40/

39 *Computation and Human Experience*, op. cit.

40 See above, Brooks's subsumption architecture.

In Agre's work, metaphorical analysis can become the basis for widening our perspective on the space of possible technologies.

Varela, Thompson, and Rosch

Francisco Varela, Evan Thompson, and Eleanor Rosch do not combine AI with cultural studies. Varela is a well-known cognitive scientist (a sister discipline of AI); Thompson and Rosch are philosophers. Nevertheless, their work is closely related to syntheses of AI and cultural studies and deserves to be addressed along with them.

In *The Embodied Mind: Cognitive Science and Human Experience*, Varela, Thompson and Rosch integrate cognitive science with Buddhism, particularly in the Madhyamika tradition[41]. They do this by connecting cognitive science as the science of cognition with Buddhist meditation as a discipline of experience. Current trends in cognitive science tend to make a split between cognition and consciousness, to the point that some cognitive scientists call consciousness a mere illusion. Instead, Varela et. al. connect cognition and experience so cognitive scientists might have some idea of what their work has to do with what it means to be an actual, living, breathing human being.

Varela, Thompson, and Rosch stress that cognitive science — being the study of the mind — should be connected to our actual day-to-day experience of what it means to have a mind. What they mean here by experience is not simple existence per se but a deep and careful examination of what that existence is like and means. They believe that your work should not deny or push /p. 40-41/

[41] Varela, Thompson, and Rosch, op. cit.

aside your experience as a being in the world. Instead, that experience should be connected to and affirmed in your work. In this way, they connect with cultural critics of science like Donna Haraway and cultural theorists like Gilles Deleuze and Félix Guattari, who stress the importance of personal experience as a component of disciplinary knowledge[42].

One of the tensions that has to be resolved in any work that combines science with non-scientific disciplines (of which Buddhism is certainly one !) is the differential valuation of objectivity. Science tends to see itself as objective, generating knowledge that is independent of anyone's individual, personal experiences. Since Varela, Thompson and Rosch want to connect cognitive science as science with individual human experience, they confront this problem of subjectivity versus objectivity head-on.

Interestingly, they do this by redefining what objectivity means with respect to subjective experiences. You cannot truly claim to be objective, they say, if you ignore your most obvious evidence of some phenomenon, i.e. your personal experience of it. This is particularly true when one is studying cognition — in this frame of thought, any self-respecting study of the mind should be capable of addressing the experience of having one !

/p. 41-42/

---

42 Donna Haraway. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective". In *Simians, Cyborgs, and Women: The Re-Invention of Nature*, pages 183-201. Free Association, London, 1990.
Gilles Deleuze and Félix Guattari. "November 28, 1947: How Do You Make Yourself a Body Without Organs." In *A Thousand Plateaus: Capitalism and Schizophrenia*, Chapter 6, pages 149-166. University of Minnesota Press, Minneapolis, 1987. Translated by Brian Massumi.

Given that one of the things cognitive scientists (and, by extension, AI researchers) are or should be interested in is subjective experience, Varela, Thompson, and Rosch abandon the focus on objectivity per se. But they stress that this does not lead to the nihilistic abandonment of any kind of judgments of knowledge which seems to haunt the nightmares of many participants in the Science Wars. Rather, they argue that Buddhist traditions have *disciplined* ways of thinking about that experience. The problem, they say, is not with subjectivity, but with being undisciplined. The goal, then, is being able to generate a kind of cognitive science that is subjective *without being arbitrary*.

Summary: Perspectives on Integrating AI and the Humanities

Generally, each of these researchers is interested in AI because of a fascination with the nature of human experience in the world. This interest naturally leads them to the humanities, which have dealt with questions of subjective human experience for hundreds of years. These researchers have found various ways to integrate this humanist experience with the science and engineering practices of AI. With respect to the issue of integrating AI and cultural studies that is pursued here, we can sum up their perspectives as follows:

- Winograd and Flores contrast existentialist philosophy with the analytic, rationalist philosophy that underlies much AI research. They use the differences between these approaches to understand the constraints that are inherent in AI methodology. They then develop new technology that, ins-tead of being limited /p. 42-43/

by these constraints, takes advantage of them.

- Suchman analyzes current AI practices to uncover the metaphors that underlie them. These metaphors turn out to be specific to Western culture. She then asks what technology would be like if it were based on metaphors from a different culture.
- Chapman implements technology that is deeply informed by, among other things, the newly-identified metaphors of Suchman. He defends the concept that, though technology is well and good, fundamental *ideas* that are not testable in a scientific or mathematical sense are equally valuable to AI.
- Agre understands technical work as reflecting deep philosophical tensions. From this point of view, technical problems *are* philosophical problems. This means that the best progress can be made in AI by thinking simultaneously at the technical and at the philosophical levels.
- Varela, Thompson, and Rosch connect the science of human cognition with the subjective experience of human /p. 43-44/

> existence. They introduce, flesh out, and defend the idea to scientists that subjective does not necessarily mean arbitrary.

While each of these researchers went a different path in integrating cultural studies and AI, often with quite different goals and self-understandings, their approaches share common themes. They are based on the idea that humanist conceptions have concrete implications for technology, and that technology can and should be changed to reflect humanist convictions and values. Their work is not a simple incorporation of cultural studies to technical ends, but also re-form both technology and the technical research process in order to align them better with a cultural studies perspective. Technical practices and cultural studies meet as equals.

CRITICAL TECHNICAL PRACTICES IN AI TODAY

In recent years, a small but active community of researchers focusing on critical technical practices has developed in AI. Researchers draw on various strands of cultural studies and cultural critique as practiced in the art community. They share a commitment to philosophical and cultural critique of technology, and its embodiment in new technical systems, which are presented to the computer science community. Three examples give an overview; they are by no means exhaustive.

Penny

Simon Penny's approach to critical technical practices, which he terms "reflexive engineering," integrates a practice of art with robotics. Penny's artworks are technical systems which embody critiques of dominant strains of thinking in robotics in /p. 44-45/

particular and computer science in general. In his work, Penny explores the aesthetics of behavior, i.e. a new aesthetics of interactivity made possible by computational and robotic machinery. Because he is an artist, he argues he is able to more freely explore possible technologies than computer scientists, who are generally constrained to generate functionally oriented, clean, and optimized systems[43].

Petit Mal, for example, is a minimalistically engineered, whimsical, elegantly clumsy robot, which interacts physically with the audience and whose chaotic behavior elicits an enormous range of culturally-specific interpretations from its audience. The tenuous relationship between Petit Mal's simple design and the audience's complex interpretation points out the extent to which our perceptions of and judgments about technical artefacts are always already embedded in a cultural environment. "Petit Mal constitutes an Embodied Cultural Agent: an agent whose function /p. 45-46/

---

[43] For general description of Penny's systems and approach, see Simon Penny, "Agents
as Artworks and Agent Design as Artistic Practice". In Kerstin Dautenhahn, ed., Human Cognition and Social Agent Technology, Amsterdam: John Benjamins, 2000. http://www art.cfa.cmu.edu/www-penny/texts/Kerstin_Chptr_Final.html
For discussion of the difficulties of being an artist doing technical work, see Simon Penny, "Consumer Culture and the Technological Imperative: The Artist in Dataspace". In Simon Penny, ed., *Critical Issues in Electronic Media*, SUNY Press, 1995. http://www-art.cfa.cmu.edu/www-penny/texts/Artist_in_D'space.html
For an example of the cultural critique embodied in this work see Simon Penny, "Virtual Reality as the End of the Enlightenment Project". In Anderson and Loeffler, eds., *Virtual Reality Casebook*, Van Nostrand, 1994. http://www-art.cfa.cmu.edu/www-penny/texts/VR_Dia.html
For an example of a computer science result of this work see Simon Penny, Jeffrey Smith and Andre Bernhardt, "Traces: Wireless full body tracking in the CAVE," ICAT virtual reality conference, Japan, December 1999. http://www-art.cfa.cmu.edu/www-penny/texts/traces

is self reflexive, to engage the public in a consideration of the nature of agency itself"[44].

Sack

Warren Sack works in computational linguistics, or the computer analysis of human language use. Using a cultural studies perspective on language leads Sack to choose unusual problems to work on. For example, most story-understanding systems attempt to extract an objective meaning from a giving piece of text. In contrast, Sack has built a system which understands ideological bias of news story by analyzing the roles the various actors in the story play. In his most recent work, Sack has created a tool, the Conversation Map, for analyzing the large-scale conversations that take place in netnews groups, including analysis of the topics of conversation, the ways in which terms are commonly used and related, and the social networks that are built in the course of conversation. Sack's goal in building this system is to be able to understand experimentally how net-based communities and subjectivities develop[45]. /p. 46-47/

---

[44] Simon Penny. "Embodied cultural agents at the intersection of robotics, cognitive science, and interactive art". In Kerstin Dautenhahn, editor, *Socially Intelligent Agents: Papers from the 1997 Fall Symposium*, pages 103-105, Menlo Park, 1997. AAAI Press. Technical Report FS-97-02. http://www-art.cfa.cmu.edu/www-penny/texts/Embodied_Cult_Agents.html

[45] For an example of Sack's cultural studies work, see Warren Sack, "Artificial Human Nature" In *Design Issues*, Volume 13, (Summer 1997): 55-64. http://www.media.mit.edu/~wsack/design-issues.html

For analysis of ideological bias see Warren Sack, "Actor-Role Analysis: Ideology, Point of View and the News". In *Narrative Perspectives: Cognition and Emotion*, Seymour Chatman and Will Van Peer (editors), New York: SUNY Press, 2000. http://www.media.mit.edu/~wsack/actor-role-analysis.html

For news group analysis see Warren Sack, "Stories and Social Networks". In Michael Mateas and Phoebe Sengers, eds., *Proceedings of the American*

Mateas

Michael Mateas engages in an AI-based art practice, which he terms Expressive AI. The goal in his work is to synthesize the development of new AI technologies with the generation of interactive artwork. An example of Mateas's work is a system called "Terminal Time," a collaboration with media artists Paul Vanouse and Steffi Dolmike. Terminal Time automatically generates ideologically-biased historical documentaries in response to audience feedback. It uses state-of-the-art story generation technology in order to demonstrate the rigidity of ideological thinking and the manipulation of historical data in historical documentaries[46].

Mateas argues, "AI-based art is not a subfield of AI, nor affiliated with any particular technical school within AI, nor an application of AI. Rather it is a stance or viewpoint from which all of AI is reconstructed"[47]. In particular, expressive AI focuses on the expression of human authorial intention through 'intelligent' machines, rather than on the generation of autonomous intelligent processes. An explicit commitment of Expressive AI is the analysis and provision of interpretive and authorial affordances, i.e. what sorts of interpretations a technical design or methodology /p. 47-

---

Association of Artificial Intelligence Symposium on Narrative Intelligence, Cape Cod, MA, November 1999. http://www.cs.cmu.edu/~michaelm/nidocs/Sack.pdf and Warren Sack, "Discourse Diagrams: Interface Design for Very Large-Scale Conversations," *Proceedings of the Hawaii International Conference on System Sciences*, Persistent Conversations Track, Maui, HI, January 2000.

46 Michael Mateas. "Generation of Ideologically-Biased Historical Documentaries." Proceedings of the 2000 Conference of AAAI

47 Michael Mateas. "Expressive AI." ACM SigGraph 2000 Electronic Art and Animation Catalog. New Orleans, 2000.

48/ supports, and the 'knobs' or 'hooks' it provides authors in order to embody their chosen concepts in the machine.

MY APPROACH

Cultural Informatics

I call my own approach to critical technical practices cultural informatics. By this, I mean a practice of technical development that includes a deep understanding of the relationship between computer science research and broader culture. This means understanding computing as a historical, cultural phenomenon, including, for example, analysis of metaphors that shape technical approaches, discovering prejudices in the Heideggerian sense that cause us to look at problems in one way to the exclusion of others, finding unconsciously held philosophical difficulties that leak their way into technical problems. These insights are used as a basis to change underlying metaphors, prejudices, philosophy, resulting in changes in technology. Cultural informatics integrates a broad humanist perspective with concrete interventions in technology and technical practices.

The approach taken in my own work follows Varela, Thompson, and Rosch in asserting that subjective experience, which goes to the heart of what it means to humans to be alive in the world, should be an important component of AI research. I believe that one of the major limitations of current AI research — the generation of agents that are smart, useful, profitable, but not convincingly alive — stems from the traditions AI inherits from science and engineering. These traditions tend to discount subjective experience as unreliable; the experience of consciousness, in this tradition, is an illusion overlaying the actual, purely mechanistic workings of our biological silicon. /p. 48-49/

It seems to me no wonder that, if consciousness and the experience of being alive are left out of the methods of AI, the agents we build based on these methods tend to come across as shallow, stimulus-response automatons.

In the reduction of subjective experience to mechanistic explanations, AI is by no means alone. AI is part of a broader set of Western cultural traditions, such as positivist psychiatry and scientific management, which tend to devalue deep, psychological, individual, and subjective explanations in favor of broad, shallow, general, and empirically verifiable models of the human. I do not deny that these theories have their use ; but I fear that, if taken as the *only* model for truth, they leave out important parts of human experience that should not be neglected. I take this as a moral stance, but you do not need to accept this position to see and worry about the symptom of their neglect in AI : the development of agents that are debilitatingly handicapped by what could reasonably accurately, if metaphorically, be termed autism. This belief that science should be understood as one knowledge tradition among others does not imply the rejection of science ; it merely places science in the context of other, potentially — but not always actually — equally valid ways of knowing. In fact, many if not most scientists themselves understand that science cannot provide all the answers to questions that are important to human beings. This means that, as long as AI attempts to remain purely scientific, it may be leaving out things that are essential to being human.

In *Ways of Thinking : The Limits of Rational Thought and Artificial Intelligence*, for example, cognitive scientist Méró, while affirming his own scientific stance, comes to the disappointing conclusion that a scientific AI will inevitably fall short of true intelligence. /p. 49-50/

In his book *Mental Models* Johnson-Laird says:

> 'Of course there may be aspects of spirituality, morality, and imagination, that cannot be modeled in computer programs. But these faculties will remain forever inexplicable. Any scientific theory of the mind has to treat it as an automaton.' By that attitude science may turn a deaf ear to learning about a lot of interesting and existing things forever, but it cannot do otherwise: radically different reference systems cannot be mixed. (228-229)[48]

But while the integration of science and the humanities is by no means a straightforward affair, the work already undertaken in this direction by researchers in AI and other traditionally scientific disciplines suggests that Méró's pessimism does not need to be warranted. We *do* have hope of creating a kind of AI that can mix these 'radically different reference systems' to create something like a 'subjectivist' craft tradition for technology. Such a practice can address subjective experience while simultaneously respecting its inheritances from scientific traditions. I term these perhaps heterogeneous ways of building technology that include and respect subjective experience 'subjective technologies.' My work is one example of a path to subjective technology, achieved through the synthesis of AI and cultural studies, but it is by no means the only possible one. /p. 50-51/

---

48 László Méró. *Ways of Thinking: The Limits of Rational Thought and Artificial Intelligence*. World Scientific, New Jersey, 1990. Edited by Viktor Mészáros. Translated by Anna C. Gósi-Greguss.

Because of the great differences between AI and cultural studies, it is inevitable that a synthesis of them will include things unfamiliar to each discipline, and leaves out things that each discipline values. In my approach to this synthesis, I have tried to select what is to be removed and what is to be retained by maintaining two basic principles, one from AI and one from cultural studies: (1) faith in the basic value of concrete technical implementation in complementing more philosophical work, including the belief that the constraints of implementation can reveal knowledge that is difficult to derive from abstract thought; (2) respect for the complexity and richness of human and animal existence in the world, which all of our limited, human ways of knowing, both rational and nonrational, both technical and intuitive, cannot exhaust.

The Anti-Boxological Manifesto

The methodologies I use inherit many aspects from the research traditions described above. Following Winograd and Flores, I analyze the constraints that AI imposes upon itself through its use of analytic methodologies. Following Suchman, I uncover metaphors that inform current technology, and search for new metaphors that can fundamentally alter that technology. Following Chapman, I provide not just a particular technology of AI but a way of thinking about how AI can be done. Following Agre, I pursue technical and philosophical arguments as two sides of a single coin, finding that each side can inform and improve the other.

The additions I make to these approaches are based on a broad analysis of attempts to limit or circumscribe human experience. I believe that the major way in which AI and similar /p. 51-52/

sciences unintentionally drain the human life out of their objects of study is through what agent researchers Petta and Trappl satirize as 'boxology:' the desire to understand phenomena in the world as tidy black boxes with limited interaction[49]. In order to maintain the comfortable illusion that these black boxes sum up all that is important of experience, boxologists are forced to ignore or devalue whatever does not fall into the neat categories that are set up in their system. The result is a view of life that is attractively simple, but with glaring gaps, particularly in places where individual human experience contradicts the established wisdom the categories represent.

The predominant contribution to this tradition of critical technical practices which I try to make is the development of an approach to AI that is, at all levels, fundamentally anti-boxological. At each level, this is done through a contextualizing approach. My approach is based on this heuristic: "that there is no such thing as relatively independent spheres or circuits"[50]. My approach often feels unusual to technical workers because it is heavily metaphorical; I find metaphorical connections immensely helpful in casting unexpected light on technical problems. I therefore include in the mix anything that is helpful, integrating deep technical knowledge with metaphorical analysis, the /p. 52-53/

49 Paolo Petta and Robert Trappl. "Personalities for Synthetic Actors: Current Issues and Some Perspectives." In Paolo Petta and Robert Trappl, editors, *Creating Personalities for Synthetic Actors: Towards Autonomous Personality Agents*, number 1195 in Lecture Notes in Artificial Intelligence, pages 209-218. Springer Verlag, Berlin, 1997.

50 Gilles Deleuze and Félix Guattari. *Anti-Oedipus: Capitalism and Schizophrenia*. Viking Press, NY, 1977. Translated by Mark Seem. p. 4.

reading of machines[51], hermeneutics, theory of narrative, philosophy of science, psychology, animation, medicine, critiques of industrialization, and, in the happy phrasing of Hayes and friends, "God knows what else." The goal is not to observe disciplinary boundaries — or to transgress them for the sake of it — but to bring together multiple perspectives that are pertinent to answering the question, "What are the limitations in the way AI currently understands human experience, and how can those limitations be addressed in new technology?"

Concretely, some of my most recent technical work is based on a tracing out and treating of the consequences of the boxological approach current in AI. I argue that the desire to construct agents in terms of a limited number of independent black boxes leads to a form of schizophrenia, or gradual incoherence in the overall behavior of the agent as more and more of these "black boxes" are combined. This schizophrenia can be traced to atomizing methodologies AI inherits from its roots in industrial culture. The disintegration AI researchers can recognize in their agents, like that felt by the assembly line worker and institutionalized mental patient, is at least in part a result of reducing subjective experience to objective atoms, each taken out of context and therefore out of relationship to one another and to the context of research itself.

This suggests that the problems of schizophrenia can be mitigated by putting the agent back into its sociocultural context, understanding its behavior as implicated in a cycle of human interpretation, on the part of both its builder and those who interact with and judge it. This approach to AI, which sees agents not in a sociocultural vacuum but as a form of communication /p. 53-54/

---

[51] Michael S. Mahoney. "Reading a Machine". In N. Metropolis, J. Howlett, and G.-C. Rota, editors, *A History of Computing in the Twentieth Century: A Collection of Essays*, pages 3-9. Academic Press, NY, 1980.

between human beings, I term "socially situated AI" and is closely related to Mateas's Expressive AI. With this metaphor as a basis, it becomes clear that creating coherence means integrating, not the agent's internally defined code, but the way in which the agent presents itself to human users. This changes the focus in agent-building from primarily a design of the agent alone, with its subsequent interpretation as an afterthought, to including the agent's comprehensibility in the design and construction of agents from the start.

Narrative psychology suggests that agents will be maximally comprehensible as intentional beings if they are structured to provide cues for narrative. I therefore argue that agent behavior should be structured as narrative, in order to make it as easy as possible for users to make coherent sense of agent activity. I implement this narrative structure for behavior using an agent architecture, the Expressivator, that connects formerly disparate behavior into coherent narrative sequences[52].

Why should a humanist care about this development? On the basis of my experience, I believe there are several advantages to using cultural studies as a basis for a practice of AI. The first is that by actually practicing AI, the cultural critic has access to a kind of experiential knowledge of science that is difficult to get otherwise and will deepen his or her theoretical analysis. This /p. 54-55/

---

52 For details on the technical implementation, see Phoebe Sengers. "Designing Comprehensible Agents" In *The Sixteenth Annual International Joint Conference of Artificial Intelligence*. Stockholm, August 1999.
http://www.gmd.de/publications/report/0077
More details on my project as a whole can be found in Phoebe Sengers. "*Anti-Boxology: Agent Design in Cultural Context*. PhD Thesis. Carnegie Mellon University. Computer Science Department Technical Report CMU-CS-98-151, 1998. http://www.cs.cmu.edu/~phoebe/work/thesis.html

increased knowledge is expressed in two ways in my work: (1) analysis of alternative AI as a manifestation of industrial culture, and (2) analysis of the metaphorical basis of alternative AI even into the details of the technology. The second advantage is that working within AI allows cultural theorists to not only criticize its workings, but to actually see changes made in practice on the basis of those criticisms. The Expressivator reflects the cultural studies analysis in the fundamental changes it makes in how an agent is conceived and structured. This brings home at a technical level the idea that agents are not simply beings that exist independently, but have authors and audiences by which and for which they are constructed.

Finally, the most important advantage to such an approach is the potential alteration to the rhetoric of mutual assured destruction that currently seems to be prevalent in interdisciplinary exchanges between cultural studies and science. The most fundamental contribution my work tries to make toward a cease-fire in the Science Wars is in demonstrating that 'science criticism' is relevant to and can be embodied in the development of technology, so that there are grounds for the two sides to respect each other, as well as a reason for them to talk. In order to address contemporary experience, we need both sides. My hope is that my work can join other similarly motivated work on whatever side of the interdisciplinary divide to replace the Science Wars with the Science Debates, a sometimes contentious and always invigorating medley of humanist, scientific, and hybrid voices.

FIRST AI, THEN THE WORLD?: THE FUTURE OF CRITICAL TECHNICAL PRACTICES

Since the 1980's, when Philip Agre began working with the approach he would later call critical technical practices, /p. 55-56/

the climate for this work has dramatically improved. What was once a few lone voices crying out in the wilderness of AI has evolved into a small research community. At the recent Narrative Intelligence Symposium[53], critical technical practices seemed to have moved into the mainstream of AI; discussion of the details of story-generation systems flowed smoothly into analyses of narrative's function in the formation of subjectivity and the role of AI narrative systems in reinforcing or undermining dominant ideologies.

But there is no reason why critical technical practices — practices of technology-building which include a critical perspective — should be limited to the subfield of AI. In fact, complementary practices have already developed and continue developing in other parts of computer science. These critical perspectives have long played a role in the field of computer-human interaction, for instance. A nice example is Kristina Höök's work, in which she develops new tools for evaluation that analyze the pleasurable quality of the experience the system provides, rather than focusing on its efficiency[54].

In a related vein, critical technical practices, and particularly cultural informatics, may have an enormous advantage in developing *poetic technology*, technical applications which enrich human life, not by making it more efficient, but by inspiring sensations of magic and wonder. Chris Dodge's "The /p. 56-57/

---

53 Michael Mateas and Phoebe Sengers, ed. *Papers from the 1999 AAAI Symposium on Narrative Intelligence*. Technical Report FS-99-01, AAAI Press, 1999.

54 Kristina Höök, Marie Sjölinder, Anna-Lena Ereback, Per Persson. "Dealing with the Lurking Lutheran View on Interfaces: Evaluation of the Agneta and Frida System." i3 Spring Workshop on Behavior Planning for Life-Like Characters and Avatars. March, 1999.

Bed" is a beautiful example of this kind of technology: it is an environment to allow intimate connection between people who are far from one another. A pillow on the bed heats when the remote participant is there, and vibrates in time with the remote person's heartbeat; a curtain moves in time with his or her breath, and colorful shadows are projected onto it according to the tenor of conversation. The result is a feeling of connection and intimacy, made possible not by optimized functionality but by the emotionally-laden overtones of the meaning of bed, light/dark, shadows, and so on[55].

Certainly, there are still gaps in the work that has been done; in particularly, in AI there has been a heavy emphasis on semiotic, philosophical, and metaphorical analysis, which can be relatively easily "smuggled into" the rhetoric of computing, with a corresponding lack of materialist analysis and work in the political economy of computing. In addition, research in critical technical practices and cultural informatics is generally done under-the-table; research communities are organized by technical application area, not by degree of incorporation of extra-disciplinary viewpoints. If research in this area is to blossom, we will probably need our own mailing lists, workshops, conferences, journals. Coherence of the community may be threatened by the heterogeneity of technical approaches, which after all may require a technically specialized audience.

Critical technical practices are generally thought of as a way of reforming the practice of computer science. A crucial question practitioners of critical technical practices will therefore have to answer is how they understand their relationship to those /p.57-58/

[55] Chris Dodge. "The Bed: A Medium for Intimate Communication". In *Proceedings of Conference on Human Factors in Computing Systems* (CHI '97). Atlanta, March 1997. ACM Press, pp. 234-241.

outside of computer science pursuing similar projects. In particular, new media art practice is often also a critical technical practice, when artists build complex computational systems (i.e. artworks) which are informed by critical reflection on technology and its role in society. The lines between technical practice, artwork, and cultural studies are blurring, and the space between is becoming home to a new interdiscipline. Hopefully, under this pressure the traditions informing the design and development of computational systems will expand, allowing for an altogether different way of looking at technology in society, and allowing for technical artefacts that enrich human experience, rather than reducing it to a quantified, formalized, efficient, and lifeless existence.

Phoebe Sengers
Media Arts Research Studies
Institute for Media Communication
German National Computer Science Research Center