SSOAR Open Access Repository
www.ssoar.info

Naming the unnamable: socionics or the sociological turn of/to distributed artificial intelligence

Malsch, Thomas

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:
Malsch, T. (2000). *Naming the unnamable: socionics or the sociological turn of/to distributed artificial intelligence* (Research Report / Technische Universität Hamburg-Harburg, Institut für Technik und Gesellschaft, 2). Hamburg: Technische Universität Hamburg-Harburg, Institut für Technik und Gesellschaft. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-408678

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Naming the Unnamable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence[1]

Thomas Malsch
Technische Universität Hamburg-Harburg
21071 Hamburg - Germany
Tel.: 040-42878-3650
E-Mail: malsch@tu-harburg.de
preprint, January 2000

[1] My thanks to Michael Florian, Ingo Schulz-Schaeffer, Holger Braun, Peter Imhof, Kai Lorentzen and Rolf Lührs for multiple and fruitful discussions that helped to shape the ideas of this article, to five anonymous reviewers of the AA-MAS journal for helpful comments, and to Paul Morland for translating the draft version and for engaging in several cooperative retranslations.

Table of Contents

1. Introduction

The roots of socionics stretch back to the late seventies and early eighties when computer scientists on the lookout for new methods and techniques for distributed and coordinated problem-solving began to take an interest in social metaphors and human society. In the course of their explorations they made contact with some sociologists, struck up a dialogue and soon found themselves, to their astonishment, involved in unexpected and strange avenues of research (Strübing 1998) into an unknown territory outside the confines of what Th.S.Kuhn used to call the normal sciences.[2] Some years later they brought out a "white paper" on coordinated problem-solving in socio-computational systems showing the need for further research and bearing the ominous title "The Unnamable" (Bendifallah et al. 1988). We have called this area of research, which was indeed then nameless, "socionics" (Malsch et al. 1996, Malsch 1997, Malsch 1998a, Müller et al. 1998).

Socionics is a new field of research, a kind of *tertium quid* between sociology and distributed artificial intelligence (DAI). Using an approach similar to that adopted by bionics in which biological phenomena serve as a source of inspiration for new technologies, socionics seeks to address the question how to exploit models from the social world for the development of intelligent computer technologies, specifically multiagent systems (MAS). To discover the borderland between sociology and DAI means to pursue the following questions: In what exactly do the characteristics of modern society consist; what makes social systems so resilient, adaptable and innovative; how may these features of modern society be translated into intelligent computer technologies; and what is the impact of sociology-based technologies on society? This set of questions has very much in common, but is by no means identical, with DAI research or with DAI-based social simulation. Instead, socionics is essentially addressed to the conceptual apparatus by which sociologists seek to observe, describe and explain modern society and from here - and only from here - it tries to build the bridge to the multiagent systems of DAI. It is an invitation to an unusual *Gedankenexperiment* where sociologists are requested to read multiagent technology as though it were a sociological text (Woolgar 1991),[3] where computer scientists are asked to read socio-

---

2 Les Gasser, one of the co-authors of the „white paper", characterized the three roots of the new research field as follows: „By the early 1980s existing DAI research had shown that the concept of coordinated problem solving was feasible and interesting. During the same period, a number of critical studies had emerged that began to seriously investigate the micro- and mid-level social aspects of computing in organizations. Concurrently, technologies for computer supported cooperative work began to appear. It was clear to some groups of researchers that an important set of related issues was beginning to appear. These concerned the prescriptive and theoretical aspects of intelligent problem solving in aggregates (from DAI and concurrent computing) and the descriptive, analytical and practical aspects of systems that involved people and machines working together with flexible dynamic divisions of labor." (Gasser/Briot 1992: 89)

3 Instead of reading „technology as object", Woolgar has suggested to read „technology as text". He argues „that the textuality of technologies and the textuality of argument is essentially similar ...that all versions (descriptions, accounts) of technology be granted no greater authority than any

logical theory as though it were a technological design and and where both groups are required to familiarize themselves with the paradox of agent societies „out of control“.

Thus, three different tenets are writ large on the agenda: (1) The first concerns the use of computer models in sociological theorizing and deals with the constitution of social order and the dynamics of social transformation. Here the claims of socionics must be substantiated in the arena of „sociological reference“. (2) The second – the lynchpin of the socionics research enterprise - is concerned with the development of new techniques and methods in DAI and investigates the role of sociological foundations in the construction of large-scale multiagent systems. Here the rules of the game are the criteria of „computational reference“. (3) The third examines the social impact of hybrid artificial societies composed of both human beings and technical agents – with possibly far-reaching consequences for our own human self-image and our very existence as social beings. And here socionics must assert itself against standards of „praxis reference“.[4] The central issue, however, is whether and how socionics will be capable of transforming sociological theories, and not just social metaphors or naive theories of sociality, into new technological potentials.

Last no least a note of caution must be made to the reader: What follows are programmatic reflections from a sociological perspective rather than research results confirmed by both disciplines; many questions are raised but remain unanswered; and where a more elaborated argumentation should be expected often only a rough outline or a tentative explication is given; and, of course, the persepective on socionics given in this paper is not the only possible one. However, introducing an unusual topic justifies a programmatic presentation of the basic ideas and the general scope of the new research enterprise.

2. DAI and Sociology

2.1 A difficult partnership

What this paper is not concerned with is a conventional exchange of methods and tools between computer science and sociology that does not touch on the central tenets of

---

other outcome of textual production and interpretation. This includes our own texts, in which we as analysts conventionally privilege our own status vis-à-vis the relativized status of the texts of others.“ (Woolgar 1991: 39ff) What we try to do with the soconics program seems to look like an operationalisation (multiagent systems as sociological text) and a rotation (sociological text as construction plan for multiagent systems) of Woolgar’s original suggestion.

4 The analogous triangulation for classic AI would be „cognitive reference“, „computational reference“ and „praxis reference“. The more current dual distinction between cognitive (basic or „strong“) AI and engineering (applied) AI appears to be less powerful - although it has helped to clarify earlier controversies about „what computers can or can’t do“ from the days of Dreyfus, McCarthy, Minsky, Searle, Simon etc. - because the concept of engineering AI is frequently misused to suggest that running („computational“) systems are useful („practical“) systems. Why computational successes often turn out to be practical failures has been shown by empirical case studies of expert systems (Malsch et al. 1993).

the two disciplines. It is much more interesting to examine how interdisciplinary cross-fertilization is achieved by producing cognitive dissonance and questioning respective basic assumptions. This is by no means to belittle the importance of conventional method transfer. On the contrary: empirical sociological research does well to plunder the methodological stock amassed by AI and exploit it in the furtherance of its own research goals. Despite the lamentably low standards of computer literacy still prevalent among sociologists, remarkable work has already been done in this area, and there are indications to suggest that we are dealing with something like a success story (Bainbridge et al. 1994; Carley 1996). Even so, a note of caution should be sounded since, like true love, the course of AI in empirical sociology never did run smoothly. We only need recall the fate of a knowledge-based system for attitudinal research in social sciences launched in the early eighties with the exaggerated claim typical of that period that it could replace a human interviewer (Baurmann/Mans 1984). Today it seems more promising to apply AI based methods to text analysis, sociometrical network analysis and statistical analysis of mass data (Hummon/Carley 1993, Chateauraynaud/Charriau 1992, Teil 1991). If used to reinforce the methodology of empirical social research, AI has a valuable contribution to make and sociologists should give it the same amount of attention and respect that they pay to statistics or to ethnographical methods.

Inversely, AI can profit from sociology by applying survey methods used in empirical social research to the design of expert systems. Although they would not serve to completely eliminate the notorious “bottleneck” of knowledge-acquisition, i.e. the problem of how to get the knowledge out of the expert’s head and into the system’s knowledge base, such social scientific methods would allow expert knowledge to be collated more reliably and to be structured more appropriately (Collins 1990, Forsythe 1993). As we gather from the proceedings of knowledge-acquisition conferences, computer scientists in this area have an open ear for methodological prompts and suggestions coming from empirical social research. Whether they act on all this good advice is of course a very different question. In the day to day business of developing intelligent application systems, most of it perishes by the wayside which just goes to show that the methods employed in social sciences are not recipes from a cookbook. In principle they are just as easy to learn as the methods of AI – but they are certainly not trivial. Applying them intelligently requires years of practical experience and a deep knowledge of the science from which they emanate. For better or worse, the methods of empirical social research always come complete with a hinterland of scientific concepts. The non-specialist “user” who is not aware of this may get more than he has bargained for and can be put off for good. Disillusionment is in store for the naive knowledge engineer who wishes to collate expert knowledge using empirical methods from the social sciences just as disappointment awaits the dilettante sociologist striking out to build a knowledge-based expert system on his own without bothering to ask a knowledge engineer for help because he has a „user-friendly“ software tool at hand.

This much should be clear: Under the harmless surface of method transfer lurk dangers of contradiction and controversy which must be addressed and resolved if we are to take socionics seriously. On the one hand there is an inherent contradiction between the theoretical claims and technological achievements of AI research and, on the other, there is a latent controversy between cognitive AI and social DAI that might eventually culminate in a change of paradigm from cognition to communication. Reducing the (D)AI phenomenon to "computer aided" empirical social research, and restricting sociological curiosity to conventional methodological questions, fades out these contradictions and „invisibilizes“ the deep techtonic tensions that give birth to a new field of research and to a new epistemic praxis. In order to fully comprehend socionics we need „triangulation“ or a research program that allows us to combine the three different perspectives of sociological, computational and praxis reference.

2.2 DAI's „problem of society construction“

Classic AI takes the human mind or the human brain as the locus in which intelligent problem-solving takes place, and accordingly seeks the technological equivalents of the cognitive skills of a single human being. DAI, however, proceeds from the assumption that complex problems can only be solved by the cooperation of many autonomously acting units, and thus is concerned with making intelligent programs cooperate with one another. This objective is most clearly apparent in the domain of multiagent systems. Multiagent systems reproduce the coordinated behavior of a number of artificial agents - software programs, endowed with autonomously controlled behavior patterns which can coordinate their actions with those of other agents with a view to solving an overarching problem. However, even if DAI makes use here of models of real social systems, its primary concern is not to investigate what makes human societies tick but rather to find general, computable principles which will allow it to overcome the technical restrictions imposed by centralized control architectures.

Even a cursory glance at the social world provides ample justification for such an approach. There is no doubt that human society disposes of a first-class stock of resources for distributed software programs and decentralized computer networks and that it would be of particular interest to develop new algorithms for massive parallel computing and complex network architectures. Despite their uncontrollability and autonomy, social systems are endowed with a high degree of resilience and fault-tolerance, as the computer jargon would have it. Unlike human individuals, biological populations or ecological systems, modern society has neither unchanging environmental boundaries nor a mechanism for natural equilibrium and stability. This makes society appear fast and slow at the same time, endowed equally with extreme versatility and ultra stability. Moreover, modern society seems to command an enormous range of capabilities for self repair and reflexive adaptivity. It is these characteristics which appear to contain the material out of which the self evolving computer networks of a globalized communication infrastructure will be built.

The “core” problem that DAI research is trying to resolve does not derive from an all encompassing vision of an artificial sociality nor is it necessarily geared to sociological concepts. It has rather to do with one of those features typical for classic AI which Carl Hewitt called the problem of „logical indeterminacy“ (Hewitt 1977): What would happen, he asked, when two “microtheories”, both equally internally consistent and thus in full accord with deductive logic, lead to contrary results? In logical terms this problem is unsolvable; as Hewitt showed, it can only be settled by recourse to “negotiation”. This is why “negotiation” was one of the first social metaphors to gain general recognition in DAI (Davis/Smith 1983). The main thrust of DAI research, then, is directed towards overcoming the limits of individual machine intelligence by making use of distributed and coordinated problem-solving techniques. Directed towards developing programs for highly complex knowledge domains, it is based on the principles of negotiating conflict and managing dissent in an intelligent way. Since such principles cannot be invented merely by „computational introspection“ alone, it is at this juncture that sociology can step in. Indeed, the DAI community has always looked beyond the confines of computer science to gain new impulses from the organizational and social sciences at work on similar issues. Thus, in its own view, DAI not only has a strong vested interest in interdisciplinary discourse but is also dedicated to the investigation of social interaction and social systems. This twofold intent is underscored in many papers and reports published by the discipline and in particular by the two documentations edited by Bond/Gasser and Huhns/Singh (Bond/Gasser 1988, Huhns/Singh 1998).

Nevertheless, there are frictions between an interdisciplinary research enterprise interested in explaining phenomena of emergent cooperation and the present level of technical development in viable multiagent systems. Within the DAI community they take the form of a play-off between two different sets of research interests. On the one hand, DAI is interested in understanding the social nature of cooperation between autonomous actors by means of programming multiagent systems as a technological end in itself. In this spirit it hold out the offer of a partnership to sociology, but even while doing so, it cannot forget its role as a rival and competitor. On the other hand, the concerns of DAI as an engineering science are centered on the development of new distributed software programs that make good the limitations of sequential processing and enable the exploitation of new software programs to tap the potential of a new generation of massive parallel computing. There is a general consensus in the DAI community that both strands belong together and that there are strong bonds binding theoretical objectives and engineering science goals. And there is also wide agreement that, with respect to these twin objectives, the DAI community is up against an unresolved “problem of society construction“ (Müller 1993) and that the concepts furnished by classic AI so far are neither sufficient to deal with the foundational problems of intelligent cooperation nor to develop innovative mechanisms on which performative distributed systems may be built. Thus the issues of how to find a solution for the "problem of society construction" and how to employ everyday social metaphors or sociological concepts to the task in hand, decidedly hang in the wind.

2.3 Social metaphors in DAI

A glance back at the beginnings of DAI will make the relevance of these questions apparent. At that time one of the pioneers of DAI, Carl Hewitt, in his groundbreaking work on "patterns of passing messages" (Hewitt 1977)[5] drew on the social psychology of G.H. Mead to propose both defining the meaning of a message by the reaction it occasioned in the addressee and constructing technological AI systems as distributed and parallel systems patterned on the model of a community of scientists who all enjoy equal rights. Published as the "scientific community metaphor" (Kornfeld/Hewitt 1981), his proposal has now become a scientific by-word and heralded the advent of social metaphors in DAI literature. Since then social metaphors in DAI have become inseparable from the idea of overcoming the limitations of the "closed world assumption" and the "microtheories" of classic AI by turning to "social cooperation" and "open systems" (Agha 1986, Hewitt 1986). It is, however, also noteworthy that social metaphors have made an impact in other branches of AI, particularly with regard to connectionism and, more recently, to Artificial Life research. In this respect, inspired by the fresh winds of the connectionist ideas which were then blowing new life into the debate, in his "Society of Mind", Minsky advocated conceiving intelligence as a phenomenon that resulted from the interconnectedness of a host of subintelligent elements (Minsky 1986). At the same time a research group headed by Holland succeeded in modeling social worlds by taking a quite different track: by researching the induction problem and developing intelligent inductive models through which the discovery of natural scientific laws could be simulated (Holland et al. 1987).

All this research has colored work in the multiagent systems domain and led to a vastly expanded use of social metaphors during the 1990s. Yet although we have given some indication of their impact, the basic question regarding the precise nature of DAI interest in social metaphors like "scientific community", "negotiation", „contract", "role expectation" etc. still remains unanswered. Social metaphors are interrogated in

[5] It is noteworthy that the interesting point of Hewitt's and Agha's „actors"- and „open systems"-approach is „patterns of message passing" and not just sending messages. Recently, Agha has made this point very clear again: „Open systems are reconfigurable and extensible: they may allow components to be dynamically replaced and components to be connected with new components while they are still executing. Complex interaction patterns arise between different components of an open system. Our contention is that to simplify the task of implementing open systems in the real world, we must be able to abstract different patterns of interaction between components. On the other hand, models of concurrency are generally based on a rather low-level execution model - namely message passing as the mechanism to support interaction between components. Unfortunately, programming using only message passing is somewhat worse than programming in assembler: sending a message is not only a jump, it may spawn concurrent activity! The goal of our research is to find a set of abstractions which enable interaction patterns between concurrent components to be represented by modular and reusable code." (Agha 1997: 2)

order to find useful social laws for software development (Shoham/Tennenholtz 1992). With its manifold interactional and organizational systems, society represents a vast pool of inspirational potential for the development of concurrent algorithms and distributed computation such as will be used in future multimedia applied software in the Internet and the World Wide Web. This is why the economic potential of multiagent technology cannot be estimated too highly. Human society seems to dispose of some of the characteristics that are needed for the global communication networks of the future, and this is what makes it so interesting and attractive for a technological undertaking such as multiagent systems. And this is why we may safely assert that present efforts to develop new technologies from an amalgam of social scientific and computer technological ideas are but a mere beginning. With its socially inspired multiagent technology, socionics seems to stand in a similar favorable position to other „combined technologies“ such as bionics, Artificial Life, neuroinformatics and bioinformatics. If this is correct, then we are likely to see a vast expansion of technologically motivated interest in social metaphors in the future.

In the meantime, the readiness to learn from models of human societies and to construct intelligent computer systems along patterns of social cooperation is not only apparent in DAI, but throughout the whole of the AI community. Today the goals are set high: developing algorithms and architectures for "Artificial Social Systems" (Castelfranchi/Werner 1994). Moreover, one of the most interesting facets is the increasing awareness in the DAI community that the development of multiagent systems involves key problems of sociological theory. Carl Hewitt and Les Gasser, in collaboration with sociologists like Gerson and Star, were among the first computer scientists to propose a cross-fertilization between the two disciplines and suggested to provide adequate computational conceptions for Sociology as well as to build DAI research on sociological foundations (Gasser et al. 1989, Gasser 1991, Hewitt 1991). Although the sociological turn of DAI was originally driven by technological considerations, it seems that we are ultimately faced by questions such as what are sociological foundations like and why is social theory appropriate to modern society, i.e. how is it capable of explaining and understanding its chosen object? Confronting these questions takes us far beyond the domain of engineering tasks. Indeed we seem to be dealing with a paradoxical "mish-mash" situation in which two disciplines - rivals and yet thoroughly interpenetrated with one another - struggle to assert their claims.

At this juncture we might ask what possible profit sociological theory can expect from technical models of artificial sociality. Again, we find ourselves caught in a difficulty: DAI semantics must be thoroughly imbued with sociology before they can make useful contributions to sociological theory and vice versa. In other words which discipline underpins which? Dealing with unavoidable paradoxical and circular arguments seems to be part of the business of pursuing socionics and there is only one way out. We should begin to explore the socionics hypothesis by drawing a clear demarcation line between the respective claims of DAI and sociology, between computational and sociological reference. Thus the claim that sociological theories can be exploited technologically must be discussed and asserted nowhere else but in the arena of the computer sciences. Equally, the claim that it is possible to gain valuable sociological insights through DAI-based social models must be put to the test nowhere else but in sociological discourse.

3. Sociological Reference: On modeling sociological theories with MAS

In what sense is it possible to model theories of modern society in the medium of multiagent technologies? Given that social phenomena can not be adequately represented by introspection, how can they be grasped by computer programs in a way that is both sociologically acceptable and non-trivial? From a standpoint of „sociological reference“ the priorities are clear: to gain a profound understanding of social phenomena in all their complexity, adequate semantic concepts and appropriate theoretical languages must first be developed before we can venture onto the computer model stage. Conversely, notions and ideas that stem from the direct experience of social life are inadequate tools for arriving at a deeper understanding of society. Experience gained by direct involvement does not lend itself to unmediated conceptualization because it usually leads to insights which disguise rather than disclose social reality. Hence, whenever computer scientists set out to develop appropriate computer models of society they need to turn to the relevant specialist discipline and that is sociology.

3.1 Neither classic AI nor Connectionism can support sociological theory construction

Interestingly enough, protosociological statements seem a kind of inevitable by-product, whenever DAI concentrates on its regular work of extracting the engineering potential from distributed computer programs. Compared to other subdisciplines of computer science, DAI methods and techniques seem to lend themselves more easily to the construction of social theories. This will become clearer when we look at two of its key concepts - agent and cooperation. In order to do so we must first examine the difference between multiagent systems and „classic“ AI and „modern“ connectionism. With its two concepts of agency and cooperation, DAI has opened a path distinct both from that of the intelligence concept of traditional symbolic AI with its problem-solving programs and from that taken by the intelligence concept employed by connectionism with its neural networks. The former approach is informed by the cognitive thesis that intelligent action is to be ascribed to the single actor alone; it claims that intelligence operates according to the irreducible principles of logical symbol processing. Connectionism, on the other hand views intelligence as the functionality of the brain and attempts to reproduce this in a subsymbolic way with the help of neural networks, whereas symbolic AI propounds a concept of intelligence that is human-like, couched in terms of the Cartesian „I“ as an epistemological concept of a human being living in splendid isolation from the rest of mankind. Or in Gasser’s words: „Current AI is largely a-social, and because of this, it has been inadequate in dealing with much human behavior and many aspects of intelligence“ (Gasser 1991: 108). At least, in contrast to this, subsymbolic connectionist AI may boast of a rudimentary concept of social interaction with its specific formalisms modeled on the neurons and synapses of the human brain. However, it too must beat a retreat when it comes to modeling the social interaction of human actors endowed with their own wills and consciousness. Intelligence, here, is a product produced exclusively by stupid or blind components. Connectionism lacks a concept for autonomous action just as the symbolic processing approach lacks a concept of sociality.

Both these approaches are inadequate for representing the cooperation of a number of intelligent autonomous actors in an appropriate manner and are unable to solve DAI's "problem of society construction". In marked contrast both to the Cartesian traditionalism of the famous "physical symbol systems hypothesis" and to the subsymbolic paradigm of neural networks, DAI's multiagent systems are patterned on a protosociological frame of reference comprising of many intelligent actors whose scope for voluntaristic action raises substantial problems of cooperation and coordination. DAI has hit on a problem that leads straight to the heart of sociological theory making. Unlike other branches of computer science, multiagent technology has the potential to raise claims which sociology has to take seriously. If sociology does heed these claims and seeks to sound out the protosociological potential of DAI, then it must be prepared to examine the multiagent systems in a way that goes beyond the methodological dispute between qualitative and quantitative sociology. In order to explain why the time-honored sociological dispute about the (in)adequacy of formal models is a non-starter vis-à-vis the challenges posed by DAI, the standard arguments for and against the formalization of sociological explanations must inevitably be rehashed once more. When that is out of the way, then the discussion can turn to where the problems really are located and inquire why social simulation has failed to take an adequate grasp of the core problems of sociological theory.

3.2 To formalize or not to formalize is not the question...

To what extent can social worlds be represented on the computer and how far do computer programs measure up to modern society as the object of social scientific investigation? The case for formalization or model building is advocated by a minority in the sociological world who hold to a scientific ideal taken from the natural sciences. As all other kinds of scientific knowledge do, sociological theories work with abstractions and simplifications – even those theories which refrain from propositional statements and spread rich historical material before us. Even interpretative sociology must produce abstractions and ignore much of the phenomenological wealth of social reality, boiling it down to a set of highly condensed interpretations. From the point of view of a formal, „natural scientific“ sociology, it is a waste of time to discuss whether or not, and to what extent, abstractions are needed. Instead, stringent formal models must be developed and tested to see if their explanations measure up to the requirements of precision, simplicity and consistency or not. If we take all this for granted and leave aside the epistemological debates of the period, say, from Quine to Maturana, then we may join forces with Bainbridge et al. who launched the following attack on all stripes of non-mathematical sociology: „The general public, to the extent that it has any opinion about social theory at all, probably considers it to be mere ideology. So long as theories are rambling verbal mediations punctuated with dubious metaphors, there is little defense against this accusation. (Artificial Intelligence) and mathematical formalism are compatible methods for stating theory precisely, connecting its concepts in rigorous intellectual structures, and identifying both hidden assumptions and unexpected consequences. Skillfully written simulation programs can be an excellent medium for communication of precise theoretical statements, so long as the intended audience has learned how to read programs.“ (Bainbridge et al. 1994: 431)

However, simply replacing dubious metaphors with formal definitions does not meet the requirements for elaborating powerful, reflexive sociological theories. If we wish to gain new insights into, for instance, life styles and norms, role expectations and industrial conflict, institution-building and the transformation of values, then, according to the prevalent sociological view, it is not enough to be precise, simple and rigorous. We need recourse to highly complex theory architectures and open descriptive languages of the kind that sociology has developed in the hundred years of its history. Would they then allow us to dispense with the need for differential equations or logic programs? Although many interpretative sociologists have no compunctions about using computer programs to process their empirical data, they still remain skeptical about formal models. Their main objections are: (1) Social theories feed on richly textured descriptive languages and cannot be adequately presented as decontextualized formalisms; (2) Computer models contain an objectivist bias and are not suitable for grasping concepts of self reference and reflexivity inherent in modern human society; (3) Computer programs are simplistic and therefore not capable of realistically simulating the complexity of real social worlds. Collins who, from the perspective of interpretative sociology, has conducted particularly exhaustive studies of AI and knowledge formalization in expert systems, sums up these objections in a nutshell. In his view, modelization or formalization is possible only where society itself has created highly standardized and rationalized structures of social action like the abacus or the assembly line. In general, however, as he notes „...neither regular science nor machines can model social life.“ (Collins 1992: 730)[6]

[6] The theoretical question of how computer programs could be used as a medium to address and explicate sociological problems in an interesting new way (‘sociological reference’) must be strictly separated from the practical issues of how to substitute (or reinforce) human society (organizations, institutions) with artificial social systems (‘praxis reference’). In his polemic against what he calls „models of social life“, Harry Collins obviously confounds these two aspects: „Leigh Star (1989) discovered in her collaborative work with computer scientists that sociological metaphors were adopted wholesale. Adopting a metaphor allows the difficult parts of the problem to be ignored while the terminology makes it seem as though social life is being encapsulated in programs. (...) Nigel Gilbert describes some work intended to model the growth of human societies using networked programs to model collections of individuals. (...) It seems unlikely that the computer instantiated ‘individuals’ in these programs bear much resemblance to members of soci-

With the exception of some mathematical sociologists, Collins' views would doubtless find broad agreement with most sociologists. In any case, among the more famous names shaping the course of contemporary sociology there is hardly anybody, with the exception perhaps of James Coleman, who is prepared to put in a good word for a mathematics- or algorithm-based modelization or formalization of social theory or who seriously expects substantial gains from such a move. And, with regard to Coleman's theory of rational choice, it was not so much the mathematical form it was cast in that attracted attention as its conceptual semantics and theoretical architecture. We should not forget that the main point here is the semantic content of the theory, not the degree of its formalization. From this perspective, questions as to whether we should stick to the salutary path of „lean formalization" or prefer „thick description" are of secondary importance. Ultimately, the choice we are faced with is not that between rigid formalism and tentative description, but rather that of selecting between a number of different conceptual frameworks tailored to respective goals of research. Thus the degree of formalism to be used should depend on the nature of the problems and themes being dealt with.

3.3 ...but addressing Sociology's key theory problems

We should not be worrying whether abstraction and model building are activities which can or should be employed but should rather direct our attention to two key questions: first, can a model or a theory address sociology's key theory problems in a promising new way; and second, is the model seriously debated by the sociological community and widely accepted as an interesting contribution to resolving these problems. This is to place high demands not only on the gains expected from abstraction, but in particular on the conceptual content of sociological theorizing. It is unacceptable to build multiagent systems using some kind of intuitively grounded protosociological concept and then proclaim them sociologically interesting models of artificial sociality just because they are neatly composed, rigorously structured and hence programmable. What we should rather be seeking to produce is theories that are grounded in central sociological issues and research questions. And we can only do this after a critical reconstruction of sociology's substantive puzzles and paradoxes. This task must be conducted before DAI can even think about providing "...the social sciences with conceptual and experimental tools, namely the capacity to model and make up in parallel, reactive and cognitive systems, and the means to observe their interactions and emerging effects." (Conte/Castelfranchi 1995: v)

In order to be in a position to provide the social sciences with conceptual and experimental tools, DAI must first take on board the basic issues that have occupied re-

ety; it is more likely that they are rational abstractions of the sort used by economists and rationalist philosophers of science" (Collins 1995: 293). However, „social life" is no more „encapsulated" in Star's or Gilbert's programs than it is encapsulated in a sociological text, of course, and neither of these programs bears any more resemblance to real groups of real human beings than Collins' „core set" of scientists in his own sociological texts.

search in sociology from its very origins onwards. These issues touch on the question of just what makes sociology so distinct from its sister disciplines, yet binds its numerous currents within the framework of a permanent controversy. It is these basic problems which delineate the frame of reference within which computer-based social models must be validated and which lie at the heart of what we call „sociological reference". With all due circumspection, the core set of basic issues in sociological theory may be summarized as four succinct questions: (1) How is social order possible in the face of vigorously autonomous human beings? (2) How are social phenomena at the microlevel interrelated with the macrolevel of society? (3) How can we understand the relationship between social action and social structure? (4) How can we explain the dynamics of social transformation and structural change? It is within the arena concomitant to these questions that computer-based tools of social simulation must find their bearings and prove their worth. If they do so, then the other dispute about formal or interpretive explanations will lose much of its edge. If we accept that the construction of formal explanatory models can be as legitimate and revelatory as rich empirical descriptions or critical *Zeitdiagnoses*, then we are free at last to turn to the really important questions, namely how formal models in general or agent-based social simulation in particular can contribute to resolving the „core set problems" addressed by the micro-macro link and the emergence and transformation of social structure.

3.3.1 The micro-macro-problem: relating social agency and social structure

As we have already noted, there are different ways of posing the core set problems and each way elicits a different set of answers. Can social structures be understood as aggregates of social actions (Garfinkel); or is the reverse true and individual action largely determined by irreducible social structures (Durkheim); or do agency and structure stand in a complimentary relationship, that structures both constrain and enable social action (Giddens)? With regard to sociological theory construction, these questions are of key import for research and model building in artificial sociality. Moreover, they are also of major importance for DAI in that they indicate how agent-based (e.g. Cohen/Levesque, Wooldridge) or structure-based approaches (e.g. Hewitt, Shoham) may be integrated on the conceptual level. Here we find proposals complementary to those discussed in sociology, with the spectrum ranging from the microtheoretical approaches adopted by interpretative sociologists like Goffman and Garfinkel or game theorists like Homans and Coleman, over praxis theorists like Giddens and Habermas to systems theorists like Parsons and Luhmann. Consequently, each answer to the vexed questions of the nature of modern society furnishes the basis for the development of highly disparate formal models of artificial sociality. What is of interest to sociology here is that multiagent systems open new avenues for a methodologically controlled comparison of theories via computer simulation.

So far computer programs of social macro simulation have been unable to deliver such a comparison. The reason for this is not that they are computer programs but that they have been built to ignore the core set problems of sociological theory. Firmly rooted in the objectivist tradition of a „parameter-sociology", this kind of social simulation confines itself to describing dependency relations between variables whilst ignoring the really difficult questions our discipline must address. Of modest ambition, parameter-sociologists content themselves with being precise, simple, consistent, rigorous and so on. Rather than addressing the difficult questions of agency and social structure, traditional macro simulation deals with relationships between macro variables like age, income, gender, religious affiliation, political preference etc., using computer simulations to test the effects of parametric manipulation. The results are not uninteresting, of course, but there is no doubt that computer simulation, as practiced by conventional parameter-sociology, is not very inventive and has so far fallen miserably short of explaining sociology's basic puzzles. In fact, it is impervious to the impact of individual or collective action on social structures just as it cannot explain how structural influences impinge on individual action.

In marked contrast, micro simulations (Manhart 1995) are a great deal more promising from a sociological „core set" standpoint. Informed by the theory of rational choice imported from micro economics, they draw on the assumption that human individuals in social contexts act strategically with a view to their own best interests. Thus individual behavior may be described by general rules of rational action which are consciously followed by individuals to achieve particular goals. Micro simulation is patterned on the theorem of methodological individualism in which all statements about social affairs may be traced back to statements about individuals. The interesting point here is that the theory of rational choice claims to be able to demonstrate how general social phenomena such as social norms or the restratification of demographic structures emerge from rational action. For instance, it claims to show, theoretically and via simulations programs, how, under the influence of individual neighborhood preferences, ethnically mixed communities or city districts gradually evolve into ethnically homogeneous areas. Admittedly, what this particular type of micro sociology cannot show is how such preferences arise, since it must systematically fade out the impact the macro level of society has on individual behavior (Blau 1987). What follows from this is that „the systematic use of computer programs in sociological research has to be combined with the question of which way sociological theory can and must be changed in accordance with the new possibilities simulations offer for sociological research. 'Computational sociologists'... seldom give answers to this question; they are mostly content if and when they are able to show that they can capture some special social processes within their simulation programs... as partisans of rational choice theory... That is of course something not to be underestimated, but it is not enough" (Klüver 1998: 15f).[7]

[7] This is neither to deny that rational choice (RC) and game theory are hegemonial in economics nor to ignore the fact that game theory has been, and still is, rather influential in DAI (e.g. coalition systems). However, here we are not dealing with the social sciences in general but with sociology. And, as one of the anonymous reviewers of this paper has remarked, „modeling the interactionist version of commitment is somewhat more tricky than an interest based concept of individual action. That is exactly the reason why RC-theories are more ‚sexy' to most DAI researchers

3.3.2 Crossing the micro-macro-bridge

To overcome the avowed limitations of methodological individualism we must cast around for other explanatory concepts. And there is no lack of candidates in this field. Among others, there are at least two outstanding theories which need to be examined: Gidden's theory of structuration and Luhmann's theory of autopoietic systems. This is not the place to follow up these theories, but what can be shown is that, after the pioneering work of Hewitt and Gasser, the micro-macro link is increasingly regarded as a crucial issue of research, both in DAI and social simulation. Cristiano Castelfranchi and Rosaria Conte, for instance, underline the need for closing the conceptual gap between micro and macro approaches (Conte/Castelfranchi 1995).[8] In their view, DAI with its multiagent technologies will enable sociology to redesign the use of algorithmic models and simulation techniques. And their view finds strong support in other authors: Multiagent systems "seem more able to 'mirror' societies and groups of people than their alternatives" (Doran/Gilbert 1994: 10), i.e. conventional mathematical simulation; and they offer a much more promising alternative to traditional simulation because they promise to "cross the micro-macro bridge" (Drogoul/Ferber 1994: 6). With multiagent systems, for the first time a simulation technique seems to be available that allows both to model the structural characteristics of social systems as the emerging result of social interaction and to model social action as shaped by social structures. This should allow to represent social phenomena of far greater complexity than was possible with traditional models and to explore a wide range of sociological concepts through simulation without prejudicing prior theoretical decisions regarding the micro-macro link. In fact, conference publications in social and organizational simulation over the last few years (edited by Gilbert/Doran 1994, Gilbert/Conte 1995, Carley 1996, Troitzsch et al. 1996, Ahrweiler/Gilbert 1998, Prietula/Carley/Gasser 1998) give cause for some optimism, although most of the contributing authors do not seem to be aware of, or interested in addressing, the gap between social simulation and sociological theory.

3.4 Methodological recommendations

Having said this, we should recall that if we wish to make full use of the sociologically interesting potential of multiagent systems, we must first cast off the blinkers imposed by conventional parameter sociology and immerse ourselves in the techniques of the-

---

than other approaches." Although I am not inclined to disagree, I should point out to recent developments in RC-based sociological approaches which seem to be much more sophisticated than what is generally presented by mainstram economics

8 Referring to the volume of sociological essays edited by Alexander et al. (1987), they note: "AI is required in the treatment of the well-known problem of the micro-macro link. Only by representing (either formally or experimentally) agents' internal mechanisms, interactions and global functions can we have a chance to solve this problem in a non-speculative way. On the other hand, in order for AI to provide a significant contribution, and deal with issues of social theory in a non-naive way, it must be able to handle social science's typical puzzles and relevant data, and closely approach existing theories." (Conte/Castelfranchi 1995: vi)

ory construction used by modern sociological theories. What we need to do is to push the claims of sociology as an intellectual discipline into multiagent systems research and to reformulate our demands from the perspective of the core set of our own unresolved theory questions. Constructing artificial social systems to serve genuine sociological ends confronts us with a problem of bilateral translation: translating sociological theory into multiagent models and re-translating these models back into the discourse of sociology.

3.4.1 „Socionic Ethnography“ between curiosity and resistance

First, in taking one or two of the major approaches[9] to the „core set“ problems as a point of departure, we should explore how to translate them into formal models without loss of authenticity, how to prevent them from being trivialized, and how to bring their innate paradoxes and puzzles into sharper focus. And in translating sociological concepts, every single step at formalization should be subject to carefully documented critique. What is needed here is a kind of “socionic ethnography”[10] that elucidates the prospects and pitfalls of formal modelization as an epistemic practice balanced between sociological resistance and curiosity, and that encourages creative deviance as well as subjecting socionic translations to hostile interrogation. This does not mean that DAI engineers are expected to translate sociological concepts into design while sociologists are expected to resist “bad” translations of their concepts. To prevent the two communities from working along two parallel tracks and to insure cross-fertilization we need to organize “tandem projects” of computer scientists and sociologists cooperating as partners and making use of a certain amount of job rotation. In doing so we are hopefully able to assert both, sociological curiosity in computational modeling and resistance against succumbing to the temptations of a “computability light”. And then perhaps we will find that socionic explorations can indeed create valuable new sociological insights.

3.4.2 Retranslating computational models into sociological debates

Second, whether or not the sociological community accepts an unusual suggestion or theoretical proposal as a valuable new insight, has nothing to do with computer programs. Here, computability is not the benchmark of „validity“. In contrast to the con-

---

[9] To name but a few of those that have already been considered for socionic translations: Pierre Bourdieu (by Florian 1998), Anthony Giddens (by Rammert 1998, Conte/Castelfranchi 1995) and Niklas Luhmann (by Bachmann 1998, Ellrich/Funken 1998, Klüver 1998).

[10] This term was coined after discussions with Holger Braun and Rolf Lührs. It seems to be quite in line with Woolgar’s „reflexivity project“ (cf. footnote 3) as well as with Amann and Hirschauer who, at least tentatively, suggest that sociological theory should be ethnographically explored as a distinctive cultural practice (Amann/Hirschauer 1997: 40). Nevertheless, unlike ethnographic studies, socionic translations do not necessarily have to „entertain a ‘parasitic’ relationship to sociological theories - i.e. an expropriative, disloyal or even destructive one - ...because“, as the two authors would have it, „the self-structuration of the fields of research cross-cuts through any type of theoretical universalism“ (Amann/Hirschauer 37f, my translation, Th.M.), since we are ultimately unable to distinguish between the „the field“ and „the author“.

ventions of conventional social simulation, the running program is not the ultimate goal of formal modelization as long as we are talking of sociological reference. The challenge is rather how to re-translate a socionic model or a program of artificial sociality into a plain sociological text that is able to speak and argue for itself and play a significant role in the sociological debates on key theory problems. Because it has nothing else to offer to sociology but the text itself - since nothing else would be accepted by the community as a sociological achievement anyway - socionics must expose itself to and be able to withstand the rigors of critical scrutiny in a theory debate in which the only means it has to convince are those of its retranslated sociological arguments.

4. Computational Reference: On reading sociological theory as technological design

Is it possible to create a new generation of intelligent technologies imaged after modern human society? This question signals a switch-over from the problem of sociological adequacy to the standpoint of computational performance. In order to examine the technological exploitability of social metaphors and sociological concepts without being troubled by the stormy debates on the proper sociological explanation of society, we must first rigorously leave "truth questions" of sociology aside. This alternation from „sociological reference" to „computer scientific reference" will then allow us to address the exciting question of the role of sociological theory in DAI as a technological enterprise. In this perspective, the only valid criteria are those of computational performance: speed and runtime, algorithmic elegance and efficiency, modularity and maintainability. Nothing more in fact than the standard criteria of computer science, these are the criteria against which the innovations produced by sociologically inspired multiagent technology must ultimately be measured. Once we have turned from sociological claims to those of computer science, the main questions that arise are whether software developed by DAI can compete with the procedures, methods and tools of other branches of computer science and whether imports from sociology can give them a cutting edge in competition. The very different question of whether computer scientists working on multiagent systems can make a significant contribution to research into modern society is completely irrelevant in the context of computational reference.

4.1 A different starting point: sociological theories instead of naive social metaphors

Nonetheless we still need to raise the matter of sociological expertise. A glance back to the beginnings of distributed problem-solving (DPS) should make it apparent why we should do so. In its early work DPS was not so much concerned with developing multiagent systems as with technical engineering work that involved (1) decomposing highly complex problems into a set of simpler subproblems and (2) bringing together the partial solutions into a coherent global solution. This work gave rise to a whole bunch of questions regarding the temporal, objective and social dimensions of cooperation and coordination - problems of a type already familiar to the social sciences.

Thus the very nature of its own engineering problems was a powerful inducement for DAI to turn to sociology. Even so, the reception of the range of conceptual approaches employed in sociology has been largely a piecemeal affair. Usually, social terms in DAI are used as mere metaphors (Castelfranchi/Werner 1994: xiii). They draw on naive and sometimes rather hazy notions about social phenomena and avoid burdening themselves with the weight of sociological concept formation. A systematic approach to sociological theorizing is more the exception than the rule. So the question that need to be raised here is what can be gained by using non-metaphoric sociological concepts, and what can DAI research learn from sociological theory?

The most courageous and compelling position is without a doubt that taken up by Les Gasser who expects that recourse to sociology will have a decisive impact on DAI by furthering its advance as a technological undertaking. According to Gasser, the conceptual problems of cooperation and coordination, decentrality and openness, asynchronity and parallelity all require sociological answers which will not be found as long as there is a fixation on the naive idea of the single agent with predetermined motives, interests and intentions. As Gasser underlines: "The traditional set of analytical categories and implementation techniques used in AI does not include fundamentally social elements; the focus is on the individual actor as the locus of reasoning and knowledge and the individual proposition as the object of truth and knowing. (...) To make substantial theoretical progress, we must first begin to lay firm social foundations for DAI research. (...) DAI systems, as they involve multiple agents, are social in character; there are properties of DAI systems which will not be derivable or representable solely on the basis of properties of their component agents. We need to begin to think through and articulate the bases of knowledge and action for DAI in the light of their social character." (Gasser 1991: 111f) This view draws on the thesis that society is not composed of individuals, in fact quite the reverse: human individuals and their minds or mental states are socially constituted and develop from social interaction.

The same holds true for the category of „commitment“ which has grown popular in DAI and which Gasser believes can only be understood and implemented as a social character along the lines of Mead's reciprocal role taking. As he comments „... the notion of commitment is distributed because the agent of commitment is a distributed entity“ (Gasser 1991: 113, comp. Castelfranchi 1995, Florian 1998). In his critique of the highly influential mentalistic approach to commitment forwarded by Cohen and Levesque, Gasser further contends that it is impossible to advance as far as social „webs of commitment“ whilst holding to an a-social concept of an agent and defining intention, in the sense of a mentalistic predetermined characteristic, as „choice with commitment“ (Cohen/Levesque 1990). Instead of this, he advocates a completely different approach: „AI research must set its foundations in ways that treat the existence and interaction of multiple actors as a fundamental category“ (Gasser 1991: 112). Only when the basic categories of individual action have been replaced by the basic categories of social interaction will it be possible to build a firm theoretical foundation for the construction of multiagent systems. In pursuit of this goal, Gasser recommends that we reconceptualize the mundane ideas of the social that have been employed in DAI so far – in other words he advocates the replacement of pre-scientific social metaphors by sociological concepts.

4.2 Scalability and dynamic adaptation: paradoxes of structuration and change

To make this clearer, let us look at the software rationale on which the design of the multiagent systems of the future is based: scalability and dynamic adaptation. At the present stage of their development, multiagent systems are still characterized by relatively simple, static models which require extensive modification before they can operate in a global network of thousands of agents.[11] Lesser cautiously seems to endorse the view that DAI is in need of a turn to sociology when he writes that agent societies „consisting of ... thousands of ... agents will need to be able to form and evolve higher order social structures ... to exploit collective efficiencies and to manage emerging situations. (...) The fundamental issue to be addressed ... is what are the basic functions and interaction patterns necessary for an agent architecture to support the construction of (adaptable) systems and to allow them to operate efficiently and robustly? The answer to this question will of necessity be speculative since there is no substantial experience in building multi-agent systems of this anticipated scale and complexity“ (Lesser 1998: 91). Indeed, one of the crucial engineering question still facing DAI research is how to build very large and highly dynamic artificial social systems which will be able to deal with the complex demands of real world applications and this involves adaptive learning in and by multiagent systems in an environment like the internet.[12] Translated into sociological language, it corresponds to problems around the micro-macro link and the dynamics of social transformation. And for DAI with a sociological perspective, this means that in its own technological self interest it will have to examine the theories on emergent social structures and social change far more close-

[11] Of course, other sociological approaches to DAI are possible, e.g social situatedness or embodiement. Such micro-approaches are not discussed in this paper because they are restricted to face-to-face interaction and co-presence of acquaintanted actors. However, if we assume that modern society largely relies on anonymous mechanisms of coordination rather than on personal contact and direct communication and that it is not based on kinship and neighbourhood, it is evident that socionics is less concerned with micro issues, unless it can be shown that these do indeed have a direct impact on macro social phenomena.

[12] With regard to the internet, Rolf Lührs suggests to by-pass the problem of scalability: since programming many thousands or millions of agents is unfeasible anyway, a more promising way to tackle the scalability problem might be to remove the load of coordination from interacting agents to the system-environment level (Lührs 1999). Accordingly, multiagent systems could either be re-described as social systems within a non-social environment (the internet as natural environment), or as social subsystems within human society (the internet as the overarching social system). The system/environment-distinction, as introduced into sociology by Parsons and Luhmann, is not a ready made solution for DAI technology. However, in line with Hewitt’s open systems semantics (OSS) it might serve as an interesting research approach to DAI.

ly than it has done so far. Once again this brings us back to the core set of sociological theory problems which – even from a quite different perspective – show themselves to be the saddle point for an interdisciplinary research program in socionics.

With regard to scalability, we will need to ascertain how communication structures are formed and stabilized by social action, thus enabling social systems of growing complexity which, in turn, provide powerful social mechanisms to coordinate and aggregate growing numbers of individual actors. This could open the way to reformulating „Open Information Systems Semantics“ or OISS (Hewitt 1991) from a sociological perspective. If large-scale artificial societies are conceived as networks operating on the principles of parallelism and openness - i.e. on asynchronous activities of permanently fluctuating participants - a unified platform for networks composed both of human and technical actors is needed that serves, in sociological terms, as an institutional framework for conflict resolution. In Hewitt's approach conflict resolution or problem solving is shaped by pre-defined semantics which are, so to speak, equivalent to a pre-existing social structure. But where is the structure and its semantics coming from? This question reveals the blind spot of pre-defined semantics which cannot show how social structures emerge from social interactions. OISS, in other words, cannot see that social structures are nothing else but routines of conflict resolution resulting from previously resolved problems or conflicts (Gasser 1991). However, any sociologically appropriate model of artificial sociality must be able to handle the „uncomputational“ paradox that its communication structures must have pre-existed before they can emerge and that they must emerge before they can come into existence. How can DAI learn from the way the paradox is treated by sociology? Of course, the difficult point here is not sociological appropriateness as such but rather decoding the sociological paradox of structuration as a guideline for the construction of technological scalability. In order for a social framework to be produced by communication or interaction – or in technical terms: in order to develop a program that is able simultaneously to solve distributed problems by communication protocols and to generate communication protocols for distributed problem-solving – it would be necessary to pre-program an institutional framework that could itself shape the intentions, convictions and goals of the individual actors and, what is of equal importance - also do the whole in reverse! (Malsch 1998b)

However, the translation of a sociological problem into an engineering question is but a starting point for further investigations. And more questions will arise: how is "personal" direct communication flanked with "impersonal" generalized media (Parsons) or capital resources (Bourdieu) of social exchange (e.g. power, money, expertise, culture) and how are these resources reinforced in personal interaction (Malsch/Schulz-Schaeffer 1997)? Or, with regard to the inner dynamics of social reproduction, how do coherent global solutions emerge from incoherent local activities (Hewitt/Inman 1991: 1411); how are stable solutions possible in the face of conflicting, concurrent, and asynchronous elementary operations? And why is this issue so crucial for DAI when a sociologist might laconically observe that modern society, far from being imbued with total global coherence, can deal equally well with global incoherence and instability? To answer all these questions, sociological concepts of „time“ and “temporalization” (Luhmann 1984, Giddens 1984) need to be merged with those of DAI (Gasser et al. 1989: 55, Bond 1990: 22) and must be systematically translated into the language of multiagent learning (Weiß/Sen 1996, Weiß 1997).

4.3 Methodological recommendations

These few remarks must suffice here to point out the direction to be taken if conceptual foundations and design principles from sociology are to be established for the building of dynamic “large-scale open systems” (Hewitt 1986). This requires a joint research program between DAI and sociology that will not succumb to the deceptive limpidity of fashionable terms like “emergence” (Ellrich/Funken 1998) or blindly try to transfer sociological concepts on a “one to one basis” into computer models. Rather, we must prepare ourselves for a laborious process of transfer and translation between the two disciplines. For obvious reasons sociology does not in itself furnish any ready made solutions but rather provides a number of different starting points. But still the skeptic computer scientist may ask: I hear the message - but is it feasible anyway? And when can we expect technologically interesting results? The answer to the first question is yes and, moreover, to the second question the answer is that its feasibility has alrady been proven. A striking example of successfully translating sociological concepts into meanwhile generally exepted and widely applied technological mechanisms in DAI are acquaintances (as in Hewitt’s Actors) and modeling other agents (as in Gasser’s MACE). And, as Carley and Gasser just recently stated again, these ideas have „now become commonplace within MAS and DAI research; however, few researchers recognize the link they are making to social theory“ (Carley/Gasser 1999: 321).[13]

However, before we can run systems of artificial sociality that open up realistic application perspectives in the Internet, we have a long and thorny way to tread. And the next steps are faced with a problem of selection. If we glance through DAI publications we may be struck by the increasing number of paeans of praise wafted in the direction of sociology. We can take this as evidence that researchers in socionics within the DAI community are now allowing themselves to be guided by a wide spectrum of different sociological theories and are no longer reliant solely on G.H. Mead and sym-

13 Another example of how new technological ideas are generated from social theory is furnished by Ishida, Gasser and Yokoo. They propose a multiagent system based on proven production systems and allowing for a novel mechanism of organizational self-design by flexibly adapting the number of active agents to the current stage of the problem solution (Ishida et al. 1992). On the one hand, here, it is obvious that the ideas of social closure and (re)opening which Gasser earlier took over from Dewey and Mead (Gasser et al. 1989) are at work once more. However, it is equally obvious that the program of organizational self-design would never have been written in the way it was without a creative switch – a break with the initial sociological concept. To explicate the paradox of structuration mentioned above, such considerations need to be examined more closely.

bolic interactionism. Instead, we find a role concept that is reminiscent of Parsons (Werner 1989), whilst other approaches are informed by Luhmann's concept of confidence (Marsh 1994, Bachmann 1998) or Gidden's theory of structuration (Conte/Castelfranchi 1995, Rammert 1998). While drawing a critical line between themselves and Gasser, Conte and Castelfranchi are correct to point out that sociology embraces highly disparate concepts of actors and action, and that in sociological discourse symbolic interactionism is but one out of many competitors (Castelfranchi/Conte 1996) even though it appears to enjoy a clear advantage so far, at least within the American DAI community (Strübing 1998). Furthermore, although Gasser explicitly refers to Mead, his own approach is not conditioned by Mead alone but also bears the imprint of Callon and Latour and their "actor-network" theory (Schulz-Schaeffer 1998).[14]

4.3.1 Evaluating technological potentials of different sociological theories

However, the task here is not to select one candidate only and ignore, by fiat as it were, all the others. Intelligent selection from the range of theory offered by sociology involves both choosing at least two different approaches that bear on the problems of scalability and dynamic adaptability and evaluating their technological potential with respect to these problems.[15] This is to qualify Gasser's assertion - without sociological foundations no substantial progress in DAI – since it opens up a range of heuristic possibilities. In view of the sheer range of paradigms offered by sociology, we should limit our selection to some of the more interesting candidates. During their transformation into formal models of artificial sociality, these candidates must be subject to a systematic comparison both with regard to technological performance and sociological adequacy. To ensure a systematic evaluation of the technological potential of sociological concepts, we must distinguish between „processes" (the practical task of building computational models) and „results" (the resulting computational models as run-

[14] For a more detailed reconstruction of Gasser's approach see my article „Acquaintance, Anonymity, Objectification" (Malsch 1998b). The crucial point is that Gasser criticizes Hewitt's predefined open systems semantics (Gasser 1991) without realizing that he uses pre-defined semantics himself in MACE (Gasser et al. 1989), where acquaintance structures are „given", while intermediate solutions are „emerging" (from unsettled to settled issues) or „demerging" (backwards from settled to unsettled issues) in the course of overall problem-solving.

[15] A methodological comparison of at least two different sociological approaches is needed not only because we want to evaluate their respective technological performances but also because we need to control the point of „over-abstraction": by cross-checking we can hopefully observe how and when the models abstracted from different sociological theories begin to converge and where they turn out identical. And if we arrive at a point of over-abstraction, where we could no longer distinguish between, for instance, social structure in Gidden's or in Luhmann's sense, or even worse, between structure and social structure, i.e. where the specific quality of a particular sociological approach or of sociological theory in general is distorted or trivialized, we could ask ourselves whether we have blundered. As we can see here, socionics is in need of a rather sophisticated methodology, of a kind of methodological opportunism, or, to put it more indulgently, of a reflexive instrumentation which combines ethnography with experimental techniques.

ning systems) and compare them with one another, processes with processes and results with results. As far as the results are concerned new socionic benchmarks are needed which, like their well known canonical counterparts in DAI (Pursuit Game, Towers of Hanoi, RoboCup), can compare technical performance levels in competing solutions. Once again these few remarks are programmatic. They address the problem of building new metrics for measuring the technological impact of sociological conceptions, but they do not demonstrate how to do it.

4.3.2 Testing the strong hypothesis of socionics by systematic comparison

As far as the comparability of different translations as process is concerned, we must ensure that the sociological concepts chosen for experimental transformation are all transformed in the same way. This is where sociological „resistance" must systematically come into play again. In order to guarantee a fair comparison between, say, two different sociological theories we must resist the temptation to trivialize or to reinterpret the underlying concepts or move away from the original ideas. To assess the technological (un)suitability or value of these two theories, the process of their translation or transformation into multiagent technology must be comparable, and comparability can only be guaranteed if the transformation process is patterned as closely as possible on the original sociological concepts. Insisting on fidelity with the initial sociological concept is not meant as a safeguard against the initial concept being violated or being "picked to pieces" like an old car wreck. The point is rather that resistance is imminent because it is the only way to methodologically control the process of translation and to guarantee a fair comparison between the resulting technology, and ultimately, the only way to „falsify" the strong hypothesis of socionics.[16]

On this basis we would not only find out what it means to read different sociological theories as technology. We could also, in comparing socionic with non-socionic solutions to the problem of dynamic scalability, gather more evidence of what the strong hypothesis of socionics can and cannot deliver. Of course, at present nobody can say for sure that the proposed grounding in sociology will turn out to have less bearing on the possible future technological success of artificial sociality than expected. And of course, a plausible case can be made for the view that technological progress in artificial sociality is driven not so much by sociological orthodoxy as by the sheer pleasure socionicists find in playing with social metaphors and algorithmic principles from DAI,

---

[16] At this point of our experimental setting, sociological and computational claims („references") appear to converge and it becomes crystal clear that a sociologist who wishes to work as an engineer must remain a fully fledged professional sociologist to do a professional job in DAI-engineering. There is no need for complicated maneuvers and fantastic inventions like an „engineer-sociologist" (Callon 1987). Instead of confusing industrial innovation with sociological theorizing, we should be very precise in what we are talking about. And in order to observe how new socionic practices are established by alternation or "boundary crossing" between sociology and DAI, we need to draw a clear distinction between the two faculties, not by presupposing an ontological *hiatus* between society and technology, but in the sense of different epistemic practices and arenas.

juggling with the conceptual foundations of sociology, picking them to pieces and recombining them with multiagent technology. Innovation in socionics, then, should be seen as an inventive process, a spontaneous “migration of social metaphors” (Malsch et al. 1996) rather than a planned transfer of concepts under the governance of sociology. From a broader perspective, however, planned transfer and spontaneous invention are two sides of the same coin. Even if sociologists cooperating with DAI adhere to the maxim of resistance against trivialization and conformity with the initial sociological concept, they still are involved in a translation process that is more akin to reworking than transferring their initial theoretical concepts. It is a process vitalized by the contradictive claims of sociologists striving to gain authentic sociological knowledge in tandem with computer scientists focussed on technological innovation.

4.3.3 Boundary crossing: observing the inventive process

What is needed are transdisciplinary research projects that allow to alternate methodologically between the claims and perspectives of sociology and DAI and to observe how new epistemic practices are established by alternation or “boundary crossing” between the two disciplines. Such projects must endow both sociologists and computer scientists with a joint responsibility; they must also allow for built-in conflict of disciplinary goals by providing frequent job rotation in order to prevent escalating conflicts and deadlocks. Socionic projects will only yield positive results and allow for creative processing if pains are taken to keep the boundaries between DAI and sociology clearly delineated so that protagonists are always sure on which side they are standing and when they are crossing over to the other side. And this is all the more important, because a demarcation line cannot be drawn and crossed by the same actor at the same time. It is impossible to observe how an unplanned migration of social metaphors occurs while busily working on a planned transfer of sociological concepts. And it is impossible to strive for sociological authenticity whilst screening sociological theories for a technologically suitable idea. Nonetheless, these are precisely the situations socionic projects are going to find themselves in and it can be easily predicted that they are going to have a great deal of trouble and excitement.

5. Praxis Reference: Artificial societies out of control?

If it is correct that sociologically informed multiagent systems have an unspecified potential of application, this is by no means to imply that this potential is also “unspecific” with regard to its practical significance and its impact on society. To address this issue we must turn from “computational reference” to “praxis reference”, i.e. from the performance criteria of pure computation to the practical criteria of applied technology and ecological validity or viability.[17] Here, the focus is on the risks and chances of ap-

[17] Referring to contingent strategies or possible projects, „viability“ is more appropriate than the current, but misleading term „validity“ which refers to an objective, given world of things. Admittedly, there are borderline phenomena juxtaposed between „computation“ and „praxis“. An outstanding example is resource management in internet-supported distributed computing: „Resource

plied artificial social systems in a democratic society. Indeed, to inquire into the practical significance of artificial sociality is to open the door on a highly contentious field of praxis: How can we imagine peaceful coexistence with an artificial society of technical agents able to act autonomously and socially competent just like human beings? Do we have to prepare for the advent of hybrid communities of artificial and human "agents" (Parunak 1996, Steiner 1996) that force us to a fundamental rethinking of the relationship between humanity and technology?

5.1 The crucial issue is not the human-machine distinction but the difference between AI's monoagents and DAI's multiagent systems

Such questions are not without precedent. In an impassioned defense of the uniqueness of human intelligence against usurpation by the "computational metaphor of mind", Hubert Dreyfus had already raised similar issues (Dreyfus 1979). However, despite the fact that Dreyfus was the moral winner of the subsequent philosophical controversy on machine intelligence, it can still be said that the computational metaphor of mind is "an extremely influential notion" (Nolan 1992) that has lost none of its purchase in AI. When sociologists joined the debate on AI it soon became apparent that there was a great deal of disagreement among them as to the nature of the distinction between (wo)man and computer. Some sociologists insisted on the social distinction of human agenthood (Collins 1990, Wolfe 1991) whilst others rubbed their hands with glee at the thought of radically deconstructing the human-machine opposition (Woolgar 1985, Schwartz 1989). Yet others tried to take the sting out of those "scenarios of gloom", in which our society is allegedly transformed into a giant computer, and began to investigate how intelligent interfaces, expert systems and translation programs were practically used and socially embedded. Their work provided a useful empirical corrective to the high flying claims of AI (Suchman 1987, Hatchuel/Weil 1990, Malsch et al. 1993, Malsch et al. 1998, Rammert et al. 1998).

Even so, it is interesting to note that, in the light of the challenges thrown down by AI, none of the company have quite succeeded in formulating and establishing an independent sociological view of the problem. None of them had any better idea than to take up the given lines of battle drawn by Dreyfus and Searle, Weizenbaum and Winograd. With hindsight we can say the reason for this failure is that they were unable to see the sociological impact of the fundamental distinction between classic AI's "monoagent" systems and DAI's multiagent systems (Malsch 1997: 4). Insofar socionics is not interested in agent modelling and architecture as such. It is rather interested

allocation in multi-agent systems is a problem that raises issues of reciprocity as well as performance and security concerns. Nodes on the world-wide web, for instance, may be willing to be part of the multi-agent system if they receive something in return for allowing foreign agents to use their resources. From the performance and security perspective, agents migrating to a node may exhibit undesirable resource consumptive behaviors, either individually or as ensembles. Similarly, network channels are a scarce resource requiring controls on how they may be used." (Jamali/Thati/Agha 1998) in modelling the interplay between agents and multiagent archtitecture. Hence, from the standpoint of socionics, the task at hand is to jumble the old lines of battle around the hegemonic power of the computational metaphor of the singular mind, and to break the hold of the gloom and doom scenarios of an inhuman society modeled on technology.

5.2 Yet another paradox: designing technology according to non-technological principles

In this sense what we need to do is to construct technology that is modeled on society. Here we touch again on the issue of transforming sociological concepts into innovations in computer science, this time, however, from a quite different perspective where we find ourselves, a bit like Alice in Wonderland, confronted with yet another curious paradox. How can we even think of designing a technology according to principles which, by their own inner nature, are of a-technological quality? Technology may be viewed as a fully controlled structure of means and ends, skillfully set to work as a cleverly devised, deterministic apparatus. Society, however, presents itself as an uncontrollable, and thus a-technical phenomenon, which, in spite of (or because of) intentional planning and steering activities, we (who?) cannot really grasp in the same way as we can control a technology "because somewhere something is getting in the way with counter-steering activities" (Luhmann 1986: 203, my translation, ThM). Hence, any attempt to seriously think an artificial society modeled on real society must be disavowed because, judged by the standards of a technology under control, it would inevitably end up either in something useless or something dangerous.

To make ourselves more familiar with the paradoxical idea of an a-technical technology, we should remember that technology as a social project can only get out of control because society itself is uncontrollable. This is certainly neither a reason to be fatalistic nor to be optimistic; it is rather an opportunity to rethink, and possibly revalue, the meaning of "uncontrollability" in a socio-technical context (Malsch 1999). As far as AI is concerned, it always took particular pride in the fact that there was no way of determining in advance how intelligent programs would react when faced with a concrete problem – or at what concrete solution they would arrive. This was valid at the time for expert systems and is even more valid now for neuronal nets. We might know and program the principles that make these systems function and we might equip them with a particular operational goal such as pattern recognition. But within this predetermined framework they operate in a non-determinist and, in principle, unpredictable manner. AI rhetoric sometimes plays down the instrumentality and transparency of its artifacts whilst taking particular pleasure in highlighting the "mysterious" non-transparency of neuronal nets running on a non-deterministic operational level (Malsch 1992: 164). If we subtract the rhetoric, however, we are still left with a notable difference to "inscrutable" large technological systems such as an airbus or an atomic power station: such large systems are supposed to operate in a deterministic manner. And it is not in spite of - but precisely because of - their deterministic operating mode that they can run out of control.

5.3 From disobedient avatars...

However, agent technology and multiagent systems do force “uncontrollability” further down the line of non-deterministic operational modes. The autonomy of action granted these artifacts implies that agents might possibly do, as the Anglican Prayer Book has it, “those things which they ought not to have done”. The so-called “avatars” and “assistant agents” are an important step in this direction even if they were originally conceived as out and out ministering angels. Intelligent technologies which possess social as well as cognitive skills, they are able to “empathize” themselves into the role of the user (Braun 1998: 192).[18] They have also raised hopes that the adaptivity and flexibility of multiagent systems will be substantially increased when users and software agents are able to directly communicate with, and learn from, one another as interaction partners. A vital prerequisite for this is that both human users and technical agents „perceive“ one another as social actors. Research work in this direction orients itself on the vision of a hybrid community of humans and artificial agents that “do not distinguish between interactions with humans and interactions with artificial agents” (Parunak 1996: 150). This appears to be leading to a qualitatively new form of embedding technology into the social context of its practical application.[19]

An assistant agent which presorts its user’s electronic mail or arranges appointments is not so much a technical tool which processes instructions as a decision-maker which uses its own judgements to facilitate the daily routine of its user and yet at the same time follows its own priorities and interferes in the user’s affairs. We may suppose that assistant agents will only be equipped with such decision-making powers when they behave towards their user in the same way as a personal secretary with whose competencies they are inscribed. This opens new perspectives for reciprocal dynamic adaptability which will allow far greater account to be taken of users’ wishes than has been possible so far with the human-machine interface (Lenzmann/Wachsmuth 1997). But what would happen if the „assistent“ began to turn against its master or if the “autonomous personal representative” of a higher ranking person appeared in a position of strength before a person of lower rank and made use of its superior resources of power? What impact would this have on social relations and under what conditions would the user of the future be prepared to accept it?

---

[18] Quoting Pickering that „we humans differ from non-humans precisely in that our actions have intentions behind them“ (Pickering 1993: 565), Braun (1998: 170) cannot resist the temptation to ask what exactly to do with the difference of „having intentions“ if machines would be programmed with intentionality as in the case of agents endowed with a so-called BDI (belief, desire, intention)-architecture?

[19] To evaluate multiagent systems in hybrid settings see Star’s (1989) suggestion of what she called the „Durkheim Test“.

These are questions concerning multiagent systems under „real world“ conditions, and they must be answered to meet the requirements of practical application: feasibility and usability, acceptability and profitability, security and sustainability. Today, agent societies are still far from meeting these criteria nor will they do so unless we seriously begin to address the deep questions concerning the insertion of social agents in the context of human sociality. To evaluate the chances and risks of hybrid communities we have to analyze how the interpenetration or amalgamation of human and technical actors may be resolved in accordance with the values of a democratic society. Here, it is not enough to mobilize the means and methods of conventional technology assessment because the case of artificial societies confronts us with an unconventional challenge. For here we are having to deal not merely with the usual questions regarding the implantation of a deterministic technology in a non-deterministic context characterized by interaction relationships between social actors. We are rather faced with the paradox of a presumptive technology which, like its social context of application, is composed of non-deterministic or contingent "social relations".

The homologous nature of agent technology and human society complicates rather than simplifies matters. This applies particularly to issues of the credibility and acceptance of assistant agents by their human co-players and opponents. Assistant agents will first have to win this approval by showing they are capable of appropriate social behavior in particular situations of human interaction. They will have to exhibit a behavior that meets the demands of our shared cultural practices, even when the field of action is restricted to the narrow subset of delegated tasks. Even then agents will not be allowed to follow a predetermined path because social situations are characterized by intersecting rationales of action which must be resolved in an acceptable way. To be accepted as co-partners, assistant agents must learn that our decision-making is seldom oriented on unambiguous preference orders. And, for our part, we shall have to learn that an autonomous personal representative on occasions is just as capable of independent (re)action as its owner. It is within this charged area that the possibilities and limits of socially acceptable artificial agents must be researched and evaluated (Schulz-Schaeffer 1998).

5.4 ...to new dimensions of uncontrollability

However, the question of a non-deterministic technology composed of contingent "social relations" has another, and clearly portentous, dimension. This should be apparent if we think for a moment of what lies, or could lie, beyond the human-machine interface and what is not visible on the screen. Beyond the interface there is not just one single personal assistant acting on orders, which must be watched over and controlled like a tiny golem because it is equipped with certain autonomous decision-making powers. The assistant agent is rather meshed in a network of a vast number of other agents which as "avatars" are representing other humans or perhaps – and this is the crux of the matter – are not representing any real person at all, not acting on anyone's orders, but fulfilling some other useful or reproductive function for the agent network in order to keep it going. In other words, they are acting on behalf of the agent society and not on behalf of any human user. They represent the agent society itself by working for its self-sufficiency or autarchy, and they are acting independently of – and in some cases even in direct opposition to – the interests and wishes of human users. It is from this perspective that the autopoietic reproduction of an artificial society needs to be dealt with in a critical yet open and unbiased manner.

Indeed the peculiar feature of this kind of artificial society lies precisely in the fact that it is as little at the command of our wills as "our" real society is. Whilst an uncontrollable assistant agent would be a deficient construction, quite the reverse holds true for an uncontrollable society of agents: trying to build artificial societies on the model of human societies means mobilizing the technological potential of an essentially non-technological modus of social reproduction which can no longer be manipulated "from the outside" – in other words by designers, users and owners. Even if an artificial society were composed exclusively of „personal assistants“ - i.e. without any additional internal agents acting on behalf of its autarchy - and even if all these personal assistant agents were endowed with as little autonomy as possible, each of them working to the strict orders of their human owner „outside“, it would still be impossible, for systematic reasons, to command and control the agent society as a whole - just as modern society, although composed of nothing less than our own self-ordained actions, lies beyond the sphere of our control. Only when we have come to terms with the idea that autonomous artificial societies can be something quite different from the industrious little dwarves of our fairy tales (Rammert 1998) or the ravaging Internet agent bands of our nightmares, will we be free to pursue the paradoxical question of an a-technical technology in all its radical creativity.

6. Towards a triangular research program

In order to make progress in socionics, we must clearly distinguish between the three different perspectives it offers and mesh them in a triangular research program: modeling sociological theories in the medium of multiagent technology ("sociological reference"); reading sociological theories as instructions how to build innovative multiagent systems („computational reference“); and designing hybrid societies of artificial agents and human actors in accordance with social sustainabilty, democratic values and economic efficiency („praxis reference“). The issue at stake here is whether socionic research can help to bridge the growing gap between hardware potentials and software applications and supply the coming generation of massive parallel computing with complex „social“ networks sustained by a host of smart mini algorithms. Even though we do not have to begin from zero, and even though the borderland between sociology and DAI is no longer an unexplored *terra incognita* as it was a few years ago, we are still very much in the dark as to what exactly will happen when sociological models are translated into the language of DAI and as to the possible benefits or drawbacks of sociological theorizing inspired by DAI. And we still know far too little about inserting sociological theories into the inventive process of building dynamic large-scale multi-agent systems.

In order to address these issues in a way beneficial for both communities, computer science and sociology alike, we will need patience and stamina. When we break into research territory outside the confines of our well established knowledge domains, we cannot tell in advance what exactly is going to happen nor insure ourselves against possible set-backs. We cannot know at present what DAI and sociology will really learn from one another nor how they will learn it. Equally, we cannot know all the implications of developing an intelligent computer technology that takes social systems as its model. In spite of all the many unanswered questions, however, we may take heart for our expeditions in the socionic borderland from the precedent of classic AI. As we know today, the famous Dartmouth Conference of 1958 which inaugurated AI research, triggered off a paradigm revolution in the computer sciences. Forty years ago, when AI was a newcomer to computer sciences, its implications were not very clear and it had to contend with opposition from traditional computer science branches. We should keep this in mind as we strike out into a research field that not so long ago was still considered “The Unnamable”.

Literature

Agha, Gul (1986): Actors: A Model of Concurrent Computation in Distributed Systems. Cambridge, Mass.: MIT Press

Agha, Gul (1997): Abstracting Interaction Patterns: A Programming Paradigm for Open Distributed Systems, in: E. Najm / J.-B. Stefani (eds.): Formal Methods for Open Object-based Distributed Systems, Chapman & Hall

Ahrweiler, Petra / Gilbert, Nigel (eds.) (1998): Computer Simulations indf Science and Technology Studies, Berlin, Heidelberg: Springer

Alexander, Jeffrey C. / Giesen, G. / Münch, R. / Smelser, N. J. (eds.) (1987): The Micro-Macro Link, Berkeley, London: University of California Press

Amann, Klaus / Hirschauer, Stefan (1987): Die Befremdung der eigenen Kultur: Ein Programm, in: S. Hirschauer / K. Amann (eds.), 7-52

Attabou, Rhazi / Chabot, Robert (eds.): Science sociales et intelligence artificiel, Technologies, idéologies, pratiques Vol. X (2-4) 1992, Aix-en-Provence

Avouris, Nicolas M. / Gasser, Les (eds.): Distributed Artificial Intelligence: Theory and Praxis, Dortrecht etc.: Kluwer

Bachmann, Reinhard (1998): Kooperation, Vertrauen und Macht in Systemen Verteilter Künstlicher Intelligenz. Eine Vorstudie zum Verhältnis von soziologischer Theorie und technischer Modellierung, in: Thomas Malsch (ed.), 197-234

Bainbridge, William S. / Brent, E. / Carley, K. / Heise, D. / Macy, M. / Markovsky, B. / Skvoretz, J. (1994): Artificial Social Intelligence, in: Annual Review of Sociology 20/1994, 407-436

Baurmann, M./Mans, D. (1984): Künstliche Intelligenz in den Sozialwissenschaften. Expertensysteme als Instrumente der Einstellungsforschung, in Analyse & Kritik 6/1984, 103-159

Bendifallah, Salah / Blanchard, F. / Cambrosio, A. / Fujimura, J. / Gasser, L. / Gerson, E.M. / Henderson, A. / Hewitt, C. / Scacchi, W. / Star, S. L. / Suchman, L. / Trigg, R. (1988): The Unnamable: A White Paper on Socio-Computational ‘Systems’, unpublished draft manuscript available from Les Gasser, Department of Computer Science, University of Southern California, Los Angeles

Bijker, Wiebe / Hughes, T. P. / Pinch, T. (eds.)(1987): The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology, Cambridge, Mass.: MIT-Press

Blau, Peter M. (1987): Contrasting Theoretical Perspectives, in: J. Alexander et al. (eds.), 71-85

Bond, Alan H. (1990): A Computational Model for Organization of Cooperating Intelligent Agents, in: Proceedings International Workshop on Office Information Systems, Cambridge, MA, 21-30

Bond, Alan H. / Gasser, Les (1988): An Analysis of Problems and Research in DAI, in: A. H. Bond / L. Gasser (eds.), 3-36

Bond, Alan H. / Gasser, Les (eds.) (1988): Readings in Distributed Artificial Intelligence, San Mateo, Ca.: Morgan Kaufmann Publishers

Braun, Holger (1998): The Role-Taking of Technology. Vom Sozialwerden der Technik, in: Thomas Malsch (ed.), 169-196

Callon, Michel (1987): Society in the Making: The Study of Technology as a Tool for Sociological Analysis, in: W. Bijker et al. (eds.), 83-103

Carley, Kathleen M. (ed.) (1996): Artificial Intelligence within Sociology. Special Issue, Sociological Methods and Research Vol. 25 (1)

Carley, Kathleen M. / Gasser, Les (1999): Computational Organitation Theory, in: G. Weiss (ed.) 1999, 299-330

Castelfranchi, Cristiano (1995): Commitments: From Individual Intentions to Groups and Organizations, in: V. Lesser (ed.) 1995, 41-48

Castelfranchi, Cristiano / Werner, Eric (eds.) (1994): Artificial Social Systems, 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW '92, S. Martino al Cimino, Italy, July, 29-31, 1992, Selected Papers, Lecture Notes in Artificial Intelligence 830, Berlin: Springer

Castelfranchi, Cristiano / Conte, Rosaria (1995): Cognitive and Social Action, London

Castelfranchi, Cristiano / Conte, Rosaria (1996): Distributed Artificial Intelligence and Social Science: Critical Issues, in: G. M. P. O'Hare / N. R. Jennings (eds.), 527-542

Chateauraynaud, Francis / Charriau, Jean-Pierre (1992): Hétérogenèse d'une machine sociologique, in: R. Attabou / R. Chabot (eds.), 337-352

Cohen, Philip R. / Levesque, Hector J. (1990): Intention is Choice with Commitment, in: Artifical Intelligence 42/1990, 213-261

Collins, Harry M. (1990): Artificial Experts. Social Knowledge and Intelligent Machines, Cambridge, Mass.: MIT Press

Collins, Harry M. (1992): H. L. Dreyfus, Forms of Life, and a Simple Test for Machine Intelligence, in: Social Studies of Science 22/1992, 726-739

Collins, Harry M. (1995): Science Studies and Machine Intelligence, in: S. Janasoff et al. (eds.), 286-301

Davis, R. / Smith, R. G. (1983): Negotiation as a Metaphor for Distributed Problem Solving, in: Artificial Intelligence 20/1983, 63-109

Dreyfus, Hubert L. (1979): What Computers Can't Do, New York: Harper and Row

Doran, Jim / Gilbert, Nigel (1994): Simulating Societies: An Introduction, in: N. Gilbert / J. Doran (ed.), 1-18

Drogoul, Alexis / Ferber, Jacques (1994): Multi-Agent Simulation as a Tool for Modeling Societies: Application to Social Differentiation in Ant Colonies, in: C. Castelfranchi / E. Werner (eds.), 3-23

Ellrich, Lutz / Funken, Christiane (1998): Problemfelder der Emergenz. Vorüberlegungen zur informatorischen Anschlußfähigkeit soziologischer Begriffe, in: Thomas Malsch (ed.), 345-393

Florian, Michael (1998): Die Agentengesellschaft als sozialer Raum. Vorschläge zur Modellierung von „Gesellschaft“ in VKI und Soziologie aus der Sicht des Habitus-Feld-Konzeptes von Pierre Bourdieu, in: Thomas Malsch (ed.), 297-344

Forsythe, Diana E. (1993): Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence, in: Social Studies of Science 23/1993, 445-477

Gasser, Les (1991): Social Conceptions of Knowledge and Action: DAI Foundations and Open Systems Semantics, in: Artificial Intelligence 47/1991, 107-138

Gasser, Les / Briot, Jean-Pierre (1992): Object-Based Concurrent Programming and Distributed Artificial intelligence, in: N.M. Avouris / L. Gasser (eds.), 81-107

Gasser, Les / Huhns, Michael, N. (eds.) (1989): Distributed Artificial Intelligence, Volume II, London: Pittman, San Mateo, Ca.: Morgan Kaufmann

Gasser, Les / Rouquette, N.F. / Hill, R.H. / Lieb, J. (1989): Representing and Using Organizational Knowledge in Distributed AI Systems, in: L. Gasser / N.M. Huhns (eds.), 55-87

Giddens, Anthony (1984): The Constitution of Society, Berkley etc.: University of California Press

Gilbert, Nigel / Doran, Jim (eds.) (1994): Simulating Societies. The Computer Simulation of Social Phenomena, London: UCL-Press

Gilbert, Nigel / Conte, Rosaria (eds.) (1995): Artificial Societies. The Computer Simulation of Social Processes, London: UCL-Press

Hatchuel, Armand / Weil, Benoit (1990): L’expert et le système. Tome 1: Rationalisation: L’ère des savoirs. Tome 2: quatre projets d’intelligence artificielle en milieu industriel. Paris: Ecole des Mines

Hewitt, Carl E. (1977): Viewing Control Structures as Patterns of Passing Messages, in: Artificial Intelligence 8/1977, 323-364

Hewitt, Carl E. (1986): Offices are Open Systems, in: ACM Transactions on Office Information Systems, Vol. 4 (3), 271-287

Hewitt, Carl E. (1991): Open Information Systems Semantics for Distributed Artificial Intelligence, in: Artificial Intelligence 47/1991, 79-106

Hewitt, Carl E. / Inman, Jeff (1991): DAI betwixt and between: From "Intelligent Agents" to Open System Sciences, in: IEEE-Transactions on Systems, Men, and Cybernetics, Vol. 21 (6), 1409-1419

Hirschauer, Stefan / Amann, Klaus (1997): Die Befremdung der eigenen Kultur. Zur ethnographischen Herausforderung soziologischer Empirie, Frankfurt/M.: Suhrkamp

Holland, J. H. / Holland, K.J. / Holyoak, R.E. / Nisbett, P.R. / Thagard, P. (1987): Induction: Processes of Inference, Learning and Discovery, Cambridge/Mass.: MIT Press

Hummon, Norman / Carley, Kathleen (1993): Social Networks: As Normal Science, in: Social Networks 15/1993, 71-106

Huhns, Michael N. / Singh, Munindar P. (1998): Readings in Agents, San Francisco: Morgan Kaufmann

Ishida, Toru / Gasser L. / Yokoo, M. (1992): Organization self-design of distributed production systems, in IEEE-Transactions on Data and Knowledge Engineering Vol. 4 (2), 123-134

Jamali, Nadeem / Thati, P. / Agha G.A. (1998): An Actor-based Architecture for Customizing and Controlling Agent Ensembles, in: IEEE Intelligent Systems, Special Issue on Agents, to appear

Janasoff, S. et al. (eds.) (1995): Handbook of Science and Technology Studies, Thousand Oaks, Ca.: Sage

Klüver, Jürgen (1998): Modelling science as an adaptive and self-organising social system: Concepts, theories and modelling tools, in: P. Ahrweiler / N. Gilbert (eds.), 15-32

Kornfeld, William A. / Hewitt, Carl E. (1981): The Scientific Community Metaphor, in: IEEE Transactions on Systems, Man and Cybernetics Vol. 11 (1), 24-33

Kuhn, Thomas S. (1962): The Structure of Scientific Revolutions, Chicago: University of Chicago Press

Lenzmann, Britta / Wachsmuth, Ipke (1997): A User-Adaptive Interface Agency for Interaction with a Virtual Environment, in: G. Weiß (ed.), 202-221

Lesser, Victor R. (1998): Reflections on the Nature of Multi-Agent Coordination and its Implications for an Agent Architecture, in: Autonomous Agents and Multi-Agent Systems, Vol. 1 (1), 89-112

Lesser, Victor R. (ed.) (1995): ICMAS-95 Proceedings. First International Conference on Multi-Agent Systems, June 12-14, San Francisco, Ca., Menlo Park: AAAI-Press

Luhmann, Niklas (1984): Soziale Systeme. Grundriss einer allgemeinen Theorie, Frankfurt/M.: Suhrkamp

Luhmann, Niklas (1986): Ökologische Kommunikation, Opladen: Westdeutscher Verlag

Lührs, Rolf (1999): Agenten, 'bots und Avatare - Überlegungen zum Verhältnis von Sozialität und Technik angesichts innovativer Softwareartefakte und ihrer Nutzung im Internet, unpublished manuscript, Technische Universität Hamburg-Harburg

Malsch, Thomas (1992): Vom schwierigen Umgang der Realität mit ihren Modellen: "Künstliche Intelligenz" zwischen Validität und Viabilität, in: Th. Malsch, U. Mill (eds.), 157-184

Malsch, Thomas (1997): Die Provokation der "Artificial Societies". Warum die Soziologie sich mit den Sozialmetaphern der Verteilten Künstlichen Intelligenz beschäftigen sollte, in: Zeitschrift für Soziologie, Vol. 26 (1), 3-22

Malsch, Thomas (ed.) (1998a): Sozionik - Soziologische Ansichten über künstliche Sozialität, Berlin: Edition Sigma

Malsch, Thomas (1998b): Bekanntschaft, Anonymisierung, Versachlichung. Skalierung von Multiagenten-Systemen als gesellschaftliche Strukturbildung, in: Th. Malsch (ed.), 255-296

Malsch, Thomas (1999): 'Artificial Societies' außer Kontrolle? Notizen zum Forschungsprogramm der Sozionik, in: Zwölfter Stuttgarter Filmwinter 14.-17.Januar 1999. Festival für expanded media, 72-73

Malsch, Thomas / Mill, Ulrich (eds.) (1992): ArBYTE. Modernisierung der Industriesoziologie? Berlin: Edition Sigma

Malsch, Thomas / Bachmann, R. / Jonas, M. / Mill, U. / Ziegler, S. (1993): Expertensysteme in der Abseitsfalle. Fallstudien aus der industriellen Praxis, Berlin: Edition Sigma

Malsch, Thomas / Florian, M. / Jonas, M. / Schulz-Schaeffer, I. (1996): Sozionik: Expeditionen ins Grenzgebiet zwischen Soziologie und Künstlicher Intelligenz, in: KI 2/1996, 6-12

Malsch, Thomas / Schulz-Schaeffer, Ingo (1997): Generalized Media of Interaction and Inter-Agent Coordination, in: Socially Intelligent Agents. Papers from the 1997 AAAI Fall Symposium, Technical Report FS-97-02, Menlo Park, Ca.: AAAI Press

Malsch, Thomas / Schwingeler, Sabine / Ziegler, Susanne (1998): Kontextualität als Orientierungsgröße für die Implementation von Expertensystemen, in: J. P. Pahl (ed.), 83-100

Manhart, Klaus (1995): KI-Modelle in den Sozialwissenschaften: logische Struktur und wissensbasierte Systeme von Balancetheorien, München: Oldenbourg

Marsh, Steven (1994): Trust in Distributed Artificial Intelligence, in: C. Castelfranchi, E. Werner (eds.): Artificial Social Systems. Proceedings, 1992. Vol. 830, Lecture-Notes in Artificial Intelligence, Berlin, Heidelberg: Springer, 94-114

Minsky, Marvin (1986): The Society of Mind, New York: Simon & Schuster

Müller, H. Jürgen (ed.) (1993): Verteilte Künstliche Intelligenz. Methoden und Anwendungen, Mannheim: BI-Wissenschaftsverlag

Müller, H. Jürgen (1993): Introduction, in: H. J. Müller (ed.), 9-21

Müller, H. Jürgen / Malsch, Th. / Schulz-Schaeffer, I. (1998): Socionics: Introduction and Potential, Journal of Artificial Societies and Social Simulation vol. 1 (3)

Nolan, J. (1992): The Computational Metaphor and Environmentalism, in: Artificial Intelligence & Society, Vol. 6 (1), 50-62

O'Hare, Greg M. P. / Jennings, Nick R. (eds.) (1996): Foundations of Distributed Artificial Intelligence, New York etc.: John Wiley

Pahl, Jörg-Peter (1998): Instandhaltung, Arbeit-Technik-Bildung, Seelze-Velber: Kallmeyer

Parunak, H. V. D. (1996): Applications of Distributed Artificial Intelligence in Industry, in: G. M. P. O'Hare / N. R. Jennings (eds.), 139-164

Pickering, Andrew (1993): The Mangle of Practice: Agency and Emergence in Sociology of Science, in: American Journal of Sociology, Vol. 99 (3), 559-589

Prietula, Michael J. / Carley, Kathleen M. / Gasser, Les (eds.) (1998): Simulating Organizations. Computational Models of Institutions and Groups, Menlo Park etc.: MIT Press

Rammert, Werner (1998): Giddens und die Gesellschaft der Heinzelmännchen. Zur Soziologie technischer Agenten und der Multi-Agenten-Systeme, in: Thomas Malsch (ed.), 91-128

Rammert, Werner / Schlese, M. / Wagner, G. / Wehner, J. / Weingarten, R. (1998): Wissensmaschinen – Soziale Konstruktion eines technischen Mediums. Das Beispiel Expertensysteme, Frankfurt/M.: Campus

Shoham, Yoav / Tennenholtz, M. (1992): On the Synthesis of Useful Social Laws for Artificial Agent Societies, in: AAAI-92. Proceedings Tenth National Conference on Artificial Intelligence, Menlo Park: AAAI Press

Schulz-Schaeffer, Ingo (1998): Akteure, Aktanten und Agenten. Konstruktive und rekonstruktive Bemühungen um die Handlungsfähigkeit von Technik, in: Thomas Malsch (ed.), 129-168

Schwartz, Ronald David (1989): Artificial intelligence as a sociological phenomenon, in: Canadian Journal of Sociology, Vol. 14 (2), 179-202

Star, S. Leigh (1989): The Structure of Ill-Structured Solutions: Boundary Objects and Heterogeneous Distributed Problem Solving, in L. Gasser / M. N. Huhns (eds.), 37-54

Steiner, Donald D. (1996): IMAGINE: In Integrated Environment for Constructing Distributed Artificial Intelligence Systems, in: G. M. P. O'Hare / N. R. Jennings (eds.), 345-364

Strübing, J. (1998): Bridging the Gap: On the collaboration between Symbolic Interactionsim and Distributed Artificial Intelligence in the field of Multi-Agent Systems Research, in: Symbolic Interaction, 21 (1998), 441-464 (also published in German: Multiagenten-Systeme als 'Going Concern'. Zur Zusammenarbeit von Informatik und Interaktionismus auf dem Gebiet der Verteilten Künstlichen Intelligenz, in: Th. Malsch (ed.), 59-90

Suchman, Lucy A. (1987): Plans and Situated Actions. The Problem of Human-Machine Interaction, Cambridge: Cambridge University Press

Teil, Geneviève (1991): Candide, un outil de sociologie assisté par ordinateur, Thèse de doctorat, Paris: Ecole des Mines

Troitzsch, Klaus G. / Mueller, U. / Gilbert, N. / Doran, J. (eds.) (1996): Social Science Microsimulation, Berlin u.a.: Springer

Weiß, Gerhard (ed.) (1997): Distributed AI meets Machine Learning, Lecture Notes in AI, Vol. 1221, Berlin: Springer

Weiß, Gerhard / Sen, S. (eds.) (1996): Adaption and Learning in Multiagent Systems, Lecture Notes in AI, Vol. 1042, Berlin: Springer

Weiß, Gerhard (ed.) (1999): Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence, Cambridge,Mass./London: MIT

Werner, Eric (1989): Cooperating Agents: A Unified Theory of Communication and Social Structure, in: L. Gasser / M. N. Huhns (eds.), 3-36

Wolfe, Alan (1991): Mind, Self, Society, and Computer: Artificial Intelligence and the Sociology of Mind, American Journal of Sociology, 5/1991, 1073-1096

Woolgar, Steve (1985): Why not a Sociology of Machines? The Case of Sociology and Artificial Intelligence, in: Sociology 19, 557-572

Woolgar, Steve (1991): The Turn to Technology in Social Studies of Science, in: Science, Technology, & Human Values 1 (16), 20-50