Social Embeddedness
– Origins, Occurrence and Opportunities

A Tutorial on "Socially Intelligent Agents" at SAB'02, at the University of Edinburgh, on the 10th August 2002, led by Kerstin Dautenhahn and Bruce Edmonds.

Introduction

It is now well established that the embedding of creatures or animats (software or robotic) in their physical environment is crucial to understanding and implementing adaptive behaviour. For social animals the social environment can be as important as the physical environment, serving many important functions, e.g.: it is an important "computational resource" which can be exploited by the individual; it can generate overwhelming complexity and uncertainty beyond the capability of the individual; and it can be the vehicle for trans-individual adaptivity. There are many parallels between physical embedding and social embedding, for example: the exploitation of social environmental phenomena as a complement to individual ability (e.g. learning from and imitating others in addition to individual learning capabilities); the use of rapid sampling of the environment instead of attempting to maintain predictive maps (e.g. gossip); and the development of regulatory systems that extend beyond the individual. However there are also important differences, including issues of internationality, agency and minds (treating others as "objects with a viewpoint/opinion/mind").

Aim of the Tutorial

The aim of the half-day tutorial is to introduce participants to some of the major themes and issues to do with Social intelligence in general and social embeddedness (SI) in particular, and social systems as it impinges on the understanding and implementation on adaptive behaviour in creatures and animats. The tutorial will continually contrast the agent and society viewpoints in an effort to get a handle on the combined individual-societal system and some of its processes.

Outline of the Tutorial

1. Introduction
2. Social Embedding - The Societal Viewpoint
   2.1. The nature of social embedding (SE)

   Analogy with embodiment and situatedness; interaction with environment; beyond one-shot & off-line interaction; web of social interaction; modelling stance towards characterising SE.

   2.2. Some of the causes of SE

   Co-evolution of social entities & abilities; parallel evolutionary processes (biological, cultural and neural); cognitive arms races; cultural adaptation to fit biological niches; development of exploitable computational and informational resources in the society.

   2.3. Some of the consequences of SE

   Impossible modelling burden for individual; importance of naming; importance of local communicative mechanisms; complexity of society and the individual; simple coping strategies (imitation, rapid sampling, games, use of proxies); emergence of new context and niches; emergence of heterogeneity.

   2.4. Some different ways understanding SE systems

   A prior vs. descriptive; bottom-up vs. top-down; different sources (philosophical, economic, ethology, ethnology, biology); different focus processes (biological processes, cognitive processes, 1-1 social interaction, social institutions and processes); different styles of model (descriptive, mathematical/logical, computational, philosophical); trans-individual entities and processes.

   2.5. Existing modelling approaches

   Economic; game theory; population dynamics; sociological theory; memetics; Alife; social robotics; social simulation; biological “models”, models from physics (e.g. self-organised criticality).

   2.6. Example: a stock market

   Imitation, arms-races, gossip and signalling, deception games, proxies, market “moods”, statistical models, chaos models, agent-based models, unpredictability, emergence of unpredictability and heterogeneity, limitations of design stance, learning, fashion.

   2.7. SE in existing social societies

   Ants; song birds, primates, humans; agents; robots; mixed societies.

   2.8. Discussion
3. Break

4. Social Embedding - Implications for the Individual and its Interactions

4.1. Phylogenetic and ontogenetic origins of social intelligence (SI)

The Social/Machiavellian Intelligence Hypothesis, social situatedness and social embeddedness, origins of human societies, the role of SI in the evolution of human intelligence

4.2. Examples of SI in humans, other primates, and other animals

Primate politics, alliances, communication and cooperation, language in non-human animals (e.g. bonobos, parrots)

4.3. Social learning and imitation in animals

Social learning mechanisms, conspecifics as social tools, definitions of imitation, agent-based perspective to imitation, imitation research in biology and psychology

4.4. Examples of imitation and social learning in robots and software

Programming by example as a new human-computer programming paradigm, imitation research in robotics, open research challenges

4.5. The relationship of SI and “theories of mind”

Mindreading, simulation-theory versus theory-theory, empathic understanding, attribution of internationality and agency, folk psychology, role of anthropomorphism in designing SIA's

4.6. SI and the origins of culture

Culture and imitation, examples of non-human culture (chimpanzees, cetaceans), implications for agent culture, culturally adaptive agents, agents that support cultural diversity

4.7. Discussion

5. Conclusion

Social Embedding[1]

Historical Introduction

As far as I am aware, Social Embeddedness was first defined in (Granovetter 1985), were he coined what he called the “embeddedness approach” to economic social phenomena. He argues against the “oversocialized approach”, were it is assumed that the individual’s behaviour is totally determined by society, and the “undersocialized approach”, where it is assumed that individuals act totally independently. He says (in the context of analysing trust and malfeasance):

> *“The embeddedness approach … threads its way between the oversocialized approach … and the undersocialized one … by following and analysing concrete patterns of social relations ... Unlike either alternative … it makes no sweeping … predictions of universal order or disorder but rather assumes that the details of social structure will determine what is found.” (page 493)*

Thus Granovetter seeks to emphasise the importance of the particular pattern of social relations an individual has.

Conceptual Introduction

In attempting to elucidate the concepts of ‘social situatedness’ or ‘social embeddedness’, one faces the problem of where to base one’s discussion. In sociology it is almost an assumption that the relevant agents are ultimately embedded in their society – phenomena are described at the social level and their impact on individual behaviour is sometimes considered. This is epitomised by Durkheim, when he claims that some social phenomena should be considered *entirely separately* from individual phenomena (Durkheim, 1893). Cognitive science has the opposite perspective – the individual’s behaviour and processes are primitive and the social phenomena may emerge as an emergent *result* of such individuals interacting.

This split is mirrored in the world of computational agents. In traditional AI it is the individual agent’s mental processes and behaviour that are the focus of their models and this has been extended to considerations of the outcomes when such autonomous agents interact. In Artificial Life and computational organisational theory the system (i.e. as a whole) is the focal point and the parts representing the agents tend to be relatively simple.

---

[1] This section is an slightly modified version of my paper (Edmonds 1999) in Adaptive Behavior.

For this reason I will take a pragmatist approach and suggest the categorisation of social systems relative to some pertinent modelling considerations. This is based on a philosophy of *pragmatic holism* which is constructivist in style. Its essence is that regardless of whether the natural world *is* theoretically reducible we have to *act* as if there are irreducible wholes. This means that we should explicitly include aspects of the modelling process in our theories. For more on this position see (Edmonds, 1996). Thus, I wish to step back from disputes as to the extent to which people (or agents) *are* socially embedded to one of the appropriateness of different types of models of agents. I want to avoid the idealisations involved in this disputed area and concentrate on what can are *useful* attributions in describing social situations and their computational analogs.

Being Situated

When Brooks (1991) made his now famous critique of AI (as it was then). He was specifically addressing shortcomings with respect to the problem of getting robots to master a *physical* environment. This spawned a whole field of research based on the premise that the physical situation was critically important in the design of agents (and in particular robots).

Since then the property of 'being situated' has been characterised in many (subtly different) ways. For example, Alonso Vera and Herbert Simon (1993) argue that the characteristics of situated action is the utilisation of external rather than internal representations via the functional modelling of the affordances provided by the environment. In their account this allows the paring down of the internal representation so that its processing can occur in real-time.

More recently William Clancey, in attempting to forge some sort of consensus on the subject wrote (page 344 of Clancy, 1997):

> *"In summary, the term situated emphasises that perceptual-motor feedback mechanisms causally relate animal cognition to the environment and action in a way that a mechanism based on logical (deductive inference alone does not capture."*

What these various approach agree upon is that if you are to effectively model certain domains of action over time then you need to include sufficient detail of the environment so that explanations of choice of action can be made in terms of the detailed causal chains in this environment. In other words, the actions will not be satisfactorily explained with reference to internal inference processes alone, but only by including causal feedback from the environment.

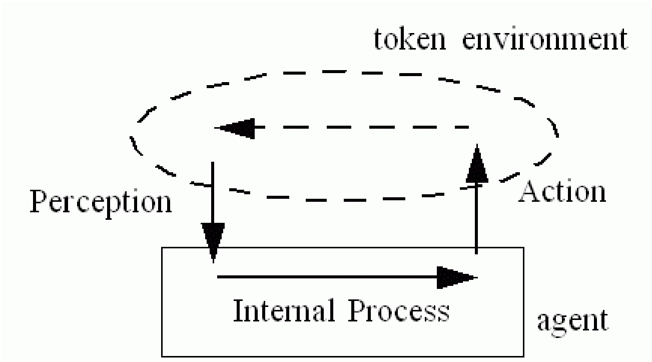

Figure 1. Where the internal inference is a sufficient as the model for action

This can be summarised (a little crudely) by saying that in a non-situated agent the internal ‘inferential’ processes form a sufficient model for the relationship between perception and action (figure 1), whereas when an agent is situated you also need to include the exterior causation (figure 2). Of course, if the agent was making a one-shot decision the pictures would be equivalent in effect since the causal part of the loop would not be needed in determining the relationship between perception and action. More usually the loop is traversed many times, with several past actions and perceptions, in order to determine the next action.

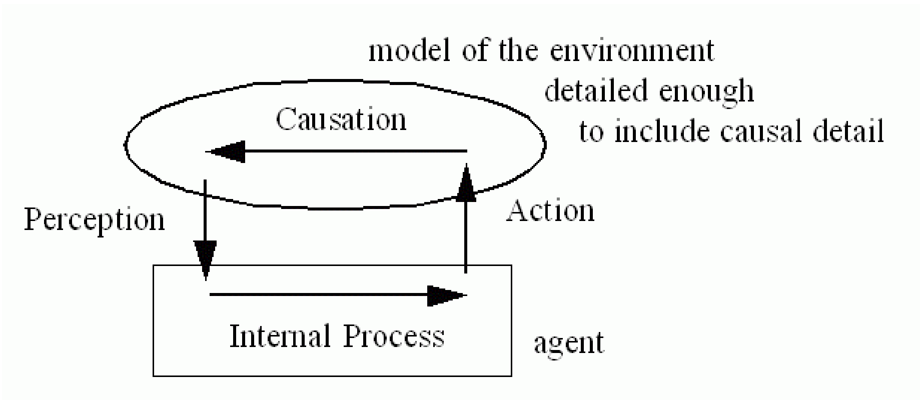

Figure 2. Where external causation is also part of the model for action

Being situated has practical considerations for what might be effective decision strategies on behalf of the agent. If internal models alone are likely to be insufficient (or just too difficult), and there are implicit computational and representational resources in the environment it make sense to make use of these by ‘probing’ them frequently for information as to effective action. This fits in with Lucy Suchman’s characterisation of situatedness which is as follows (page 179 of Suchman, 1987):

> *“... the contingence of action on a complex world ... is no longer treated as an extraneous problem with which the individual actor must contend, but rather is seen as an essential resource that makes knowledge possible and gives action its sense. ... the coherence of action is not adequately explained by either preconceived cognitive schema or institutionalised social norms. Rather the organisation of situated action is an emergent property of moment-by-moment interactions ...”*

Being Socially Situated

In a physical situation the internal models may be insufficient because of the enormous computation capacity, amount of information and speed that would be required by an agent attempting to explicitly model its environment. In a social situation, although the speed is not so critical, the complexity of that environment can be overwhelming and there is also the obvious external computational resources provided by the other agents and their interactions. This means that an agent can be said to be socially situated by analogy with being physically situated – in both cases the balance of advantage lies in using external causal processes and representations rather than internal ones. The fact that the source of this imbalance in each case is due to different causes leads to a different ‘flavour’ of the situatedness, but there is enough in common to justify the common use of word ‘situated’. Of course, social environments vary greatly and the fact of being socially situated will thus be contingent on the particular agent and its social context.

The frequent sensing and probing of the physical environment can be translated into ‘gossip’, one of whose functions is the frequent sampling and testing of the social environment. The reliance of external computational resources and models is arguably even more pronounced in social situations than physical ones – social agents may accept the output of external sources (including other agents) as a direct influence on their decision making, e.g. in fashion.

Being Socially Embedded

Extending the above characterisations of situatedness, I want to say that an agent is *socially embedded* in a collection of other agents to the extent that it is more *appropriate* to model that agent as part of the total system of agents and their interactions as opposed to modelling it as a single agent that is interacting with an essentially unitary environment. Thus saying an agent is socially embedded is stronger than saying it is merely socially situated. I have characterised social embeddedness as a *construct* which depends on one’s modelling goals, since these goals will affect the criteria for the appropriateness of models. It can be read as contrasting modelling agent interaction from an internal perspective (the thought processes, beliefs etc.) with modelling from external vantage (messages, actions, structures etc.). This is illustrated below in figure 3.

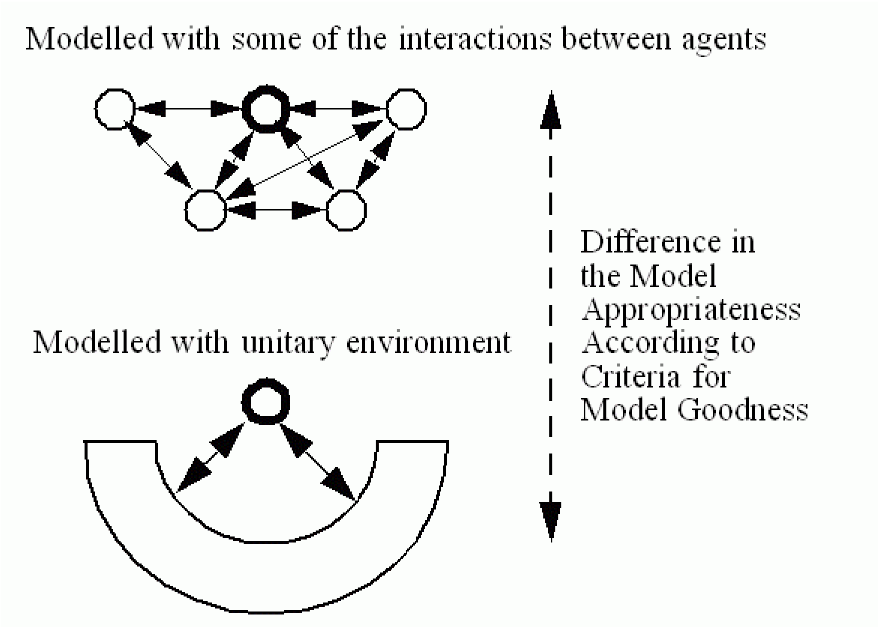

Figure 3. Social embeddedness as the appropriate level of modelling

This is not an extreme ‘relativist’ position since, if one fixes the modelling framework and criteria for model selection, the social embedding of agents within a collection of agents can sometimes be unambiguously assessed. When the modelling framework is agreed, the object of modelling (in this case ‘social systems’) will constrain the models that fit the framework. If one is extremely careful (and lucky) it might entail a unique model – in such cases we can safely project the social embeddedness upon the social system itself. Note however, that this projective attribution onto the social system is a *post-hoc* attribution that can only occur unambiguously in special circumstances. Usually there will be many arbitrary choices involved in the modelling of the social phenomena, so that the model (and hence the social embeddedness) is underdetermined by the phenomena itself. It is for this reason that it is more useful to define the social embeddedness with respect to model properties and use the association of the best model (by the chosen model selection criteria) with the phenomena itself as a means of inferring properties on the object system.

According to this account the social embedding is dependent on the modelling framework. Such a modelling framework includes the language of model representation, the model selection criteria and the goals of modelling. Frequently such a framework is implicitly agreed but not always. I have not the space here to fully specify what such a framework entails, for more details on this see (Edmonds, 1997; Moss & Edmonds, 1998).

Notice that criteria for model acceptability can include many things other than just its predictive accuracy, for example: *complexity* (Edmonds, 1997). It is the *inevitability* of these other concerns which forces us to relativise this approach as one concerning the appropriateness of our constructs (along with the different modelling goals and frameworks). For example, a computer may be able to find obscure and meaningless models which (for computational purposes) separate out the behaviour of a single agent from its society (using something like genetic programming), which are totally inaccessible to a human intelligence. Also the modelling framework is indispensable; for example, an agent may not be at all embedded from an economic perspective but very embedded from the perspective of kinship relations.

Let us consider some examples to make this a little clearer.

- *Firstly* a neo-classical economic model of interacting agents where each of these agents individually has a negligible effect on its environment, which would mean that a model of the whole system could be easily transformed into one of a representative agent interacting with an economic environment. Here one would say that each agent was not socially embedded since there is little need to model the system as a whole in order to successfully capture the agent's behaviour.
- *Secondly* where an agent which interacts with a community via a negotiation process with just a few of the other agents. Here a model which just considers an agent, its beliefs and its interaction with these few other agents will usually provide a sufficient explanation for all that occurs but there may still be some situations in which interactions and causal flows within the whole community will become significant and result in surprising local outcomes. Here one could meaningfully attribute a low level of social embeddedness.
- *Thirdly*, the behaviour of a termite. It is possible to attempt to account for the behaviour of an termite in terms of a set of internal rules in response to its environment, but in order for the account to make any *sense* to us it must be placed in the context of the whole colony. No one termite repairs a hole in one of its tunnels only the *colony* of termites (via a process of stigmergy: Grassé, 1959). Here one could say that the ants were socially *situated* but not socially *embedded*, since one can model the system with an essentially unitary model of the environment, which each of the ants separately interact with.
- *Finally*, in modelling the movements of people at a party, it is possible that to get any reasonably accurate results one would have to include explicit representations of each person and their relationship with each of the others present. This would represent a high level of social embeddedness.

At first sight this seems a strange way to proceed; why not define social embeddedness as a property of the system, so that the appropriate modelling choices fall out as a result? The constructivist approach to characterising social embedding, outlined above, results from my modelling goals. I am using artificial agents to model real social agents (humans, animals, organisations etc.). So it is not enough that the outcomes of the model are verified and the structure validated (as in Moss, Edmonds & Wallis, 1997) because I also wish to characterise the emergent process in a *meaningful* way – for it is these *processes* that are of primary interest. This contrasts with the ‘engineering approach’ where the goal is different – there one is more interested in ensuring certain specified outcomes using interacting agents. When observing or modelling social interaction this meaning is grounded in the modelling language, modelling goals and criteria for model acceptability (this is especially so for artificial societies). The validation and verification of models can not be dispensed with, since they allow one to decide which are the candidate models, but most of the meaning comes from the modelling framework. In simpler physical situations it may be possible to usefully attribute phenomena to an external reality but in social modelling we have to make too many choices in order to make progress. The proof of this particular pudding will ultimately be in the eating; whether this approach helps us obtain useful models of social agents or not.

The idea of social embedding is a special case of embedding in general – the ‘social’ bit comes from the fact we are dealing with collections of parts that are worthy of being called *agents*.

The Contribution of Society to the Construction of Individual Intelligence[2]

Introduction

Several studies of how individual intelligences can interact to allow the emergence of social structures exist. The field of Artificial Life teems with computational models composed of interacting units where it is claimed that even a modicum of 'intelligence' (in the form of some computational capacity) can result in the emergence of phenomena more usually attributed to societies. These studies are focused on the basic model that interacting units *cause* society.

What is rarer (especially outside the domains of sociology and linguistics) are investigations into the possibility that society is a causal factor in the emergence of individual practical intelligence in the individuals it is composed of.

To researchers accustomed to thinking the other way around, this may seem a little strange: for society is obviously physically composed of individuals and not the other way around. However in one supremely important respect the matter is already all but settled: humans need other humans to interact with if they are to acquire a fully functional language, and they need a fully functional language to realise much of their potential intelligence. The lack of a fully functional language does not only limit an individual's *social* intelligence, but it also limits that individual's general problem-solving ability, for example, it is inconceivable that a human without mastery of a sophisticated language could perform abstract mathematics.

The physical situation compared to the social situation

When Brooks [7] made his now famous critique of AI (as it was then). He was specifically addressing shortcomings with respect to the problem of getting robots to master a *physical* environment. This spawned a whole field of research based on the premise that the physical situation was critically important in the design of agents (and in particular robots).

Three critical aspects of being 'physically situated' are listed below, each has an analogue for the social situation.

Frequent probing and sensing – that the agent uses the frequent probing and sensing of its environment to determine its states and processes rather than attempting to use explicit models and inferential processes to predict these.

The frequent probing and sampling of the social environment of a human is called 'gossip'. We do not try and predict the details of our social environment, instead we trade information about it as frequently as we can.

---

[2] This section is a part of the paper (Edmonds and Dautenhahn 1988).

Goal directed, interactive learning – that much learning occurs in a practical interactive way in the pursuit of some specific goal rather than trying to discover general truths in a passive deductive way.

The methodology of Embodied Artificial Intelligence (EAI) approach which has influenced research into robotics and adaptive systems has according to Erich Prem [34] a number of implications for cognitive science: “Cognition is a timely process driven by forces internal and external to the system, cognition happens in close interaction with the world, often in order to manipulate the world.” [34]. If we replace ‘world’ by ‘social world’ then Prem's citation relates nicely to our notion of socially situated intelligence. For an embodied agent situatedness in the world matters, as for a social agent situatedness in the social world matters. A stronger claim, for which evidence is increasing but not yet sufficient, is that human intelligence (e.g. problem-solving abilities) has evolved in evolutionary terms literally as a side-effect of social intelligence (cf. our discussion on autism and the social intelligence hypothesis in this paper). Thus, research into socially situated intelligent, e.g. studying simulation models of human interactions/societies, or building embodied artefacts like robots, can provide valuable input to this discussion. We almost never learn about our social environment in a passive, detached way but through constant interaction with it in order to achieve our social (and other) goals.

Specific adaptations – many adaptations in a successful situated agent are very specific to the particular environment the agent occupies, exploiting its features for efficiency and effectiveness.

Humans have many adaptations that are considered as specifically social in their utility. These include: our linguistic ability; the whites of our eyes; our ability to recognise individual faces; our enjoyment of gossip; our elaborated sexuality; the expressiveness of our faces in displaying emotions; our ability to empathise with others; and our apparent predisposition towards cooperation. Thus humans are socially situated, if anything is. This does not necessarily mean that any of their features rely on this situation for its effective functioning. It may be the case that some aspects are somehow abstracted (or abstractable) from this particular situation to hold more generally. If this were the case and the abstraction preserved the feasibility then one might be able to ignore the situation and model the agent without it. On the face of it this would seem unlikely – surely the burden of proof must rest with those who would try such a task.

Why one might expect the social situation to matter

Below I outline some arguments as to why one would *expect* the social situation to be important for the development of an individual’s practical intelligence.

Language

As mentioned above, language is important for practical intelligence and language is a social construct. This means that society is critical in the development of individual intelligence unless: *either* language (once constructed) could be learnt as a entirely abstract and passive way without social interaction *or* that it might be possible to acquire a language that is not socially constructed in origin.

The first seems unlikely to be the case, at least for humans. Humans learn language in a different way as a child than later, and in fact use different areas of the brain. It seems that a first, full language with all its power can only learnt by a young child, and it is unlikely that such a child would be able to learn a language in an abstract and passive way.

The possibility of the second (an non-socially constructed language) is almost impossible to judge, because we are the only example of language users and all our (full) languages are socially constructed. It is notable that languages that are artificially devised (i.e. less socially constructed) are not as expressive or useful as full languages – on exposure to such artificial languages children seem to immediately change these into fully expressive languages in one generation as a result of their innate linguistic ability and the way they interactively acquire them (examples are the development of sign language in Nicaraguan from an artificial creation and Creole languages from pidgins).

It must be concluded that, although the last word is not in (so to speak), that language is an inherently social construct. In a related question, the possibility of private languages, there are strong arguments to say that a language can not be private [41].

The co-evolution of human intelligence and social cooperation

The recently proposed “social intelligence” [26] and “Machiavellian intelligence” theses put forward the view that substantial aspects of our intelligence evolved because its possession conferred *social* advantage. The idea is that our extensive intelligence is primarily evolved in order to keep our place in the social order and to manage the intricate cooperation and competition that this involves.

If this is indeed the case (and it would be very odd if none of our intelligent capacity has been shaped by evolutionary pressures that are socially grounded), and given the intricacy of our present society (which presupposes the possession of individual intelligence) then it seems likely that our intelligence and our society have co-evolved. If this is the case then one would expect that many aspects of our intelligence have evolved to ‘fit in’ with our society (and vice versa).

It is certainly difficult to argue from single cases, but the fact that the only species to evolve a sophisticated intelligence has also evolved a sophisticated society can not be totally ignored.

The richness of society as an informational and computational resource

One aspect of a society of roughly commensurate social agents which is almost inevitable, is that it will quickly develop so as to be more complex than any single agent can completely model. This is especially true of a society where there is sometimes advantage in 'out-guessing' the actions of the other agents, in which case a sort of modelling 'arms-race' quickly develops which in its turn makes the prediction and comprehension of the society even more difficult.

In such a complex entity it would be strange if it did not offer some informational and computational resources to some agents for certain aspects for some of the time. Given this availability it would also be odd if these resources were not exploitable by the composite agents. Hence one would expect agents that were situated in such a society to evolve ways of exploiting such social resources.

If this were the case, then we would expect that we would posses some faculties usually attributed to our 'intelligence' that were evolved to use such resources and save ourselves (individually) considerable expenditure in terms of time and effort.

The need for social reflection in the development of the self

The role of the 'self' in intelligence and consciousness is a hotly disputed one. Some philosophers see all usage of "I", "myself" and similar utterances as merely a linguistic device with no real referent [44]. Others see the self as a real entity but as one whose essence is not usefully expressible from an exterior perspective [12].

What is clearer is that: some form of self-modelling is a crucial part of the machinery of our social intelligence; this ability to model ourselves and others develops, at least partly, as the result of a learning process; this learning process requires some reflective mechanisms to occur; and that the reflection that occurs via social, linguistic mechanisms is by far the most expressive and effective that is presently available to us. In this way a reflective social environment to interact with is not only essential via language to symbolic thought but also to the development of the self.

The processes by which the self comes into existence and its relation to social reflection is unclear, but [39] makes a first cut at it and [] examines the philosophical arguments.

Autism

Since the early 40ies of this century autism is known as a syndrome which involves, among other features, the striking lack of social competence. A variety of explanations have been discussed, among them the widely discussed 'theory of mind' model which is conceiving autism as a cognitive disorder [1], and, a more recent explanation given by Hendriks-Jansen [23]. He hypothesises as the primary cause early developmental disorders which prevent the child and its caretakers to 'get the interaction dynamics right' which normally scaffold the development of appropriate social interactions in the sense of situated dialogues between infant and caretaker.

The importance of interaction dynamics are also part of the explanation given in [13] which suggests a lack of empathic processes which prevent the child to develop ‘normal’ kinds of social action and interaction. Why is it important to discuss potential explanations of the autistic syndrome? People with autism never develop into social beings as we expect of ‘normal’ people, although some of them show high intelligence in non-social domains, they are never able to communicate and interact properly with other people. They are not able to understand the social world around them, which therefore appears often scary and completely unpredictable to them. This deficit influences their lives to the extend that they often are not able to lead an independent life, in this way clearly demonstrating the central role of sociality in practical intelligence. This gives evidence that the study of socially situated intelligence does not merely provide an important add-on to other faculties of intelligence (like spatial thinking or mathematical reasoning), but that human intelligence (its development and expression) is embedded (and embodied) in a social being, and can in this way not be separated from non-social kinds of intelligence.

Consequences for the development of artificial social agents

If the above is the case and important aspects of our social intelligence require to be socially situated for their complete development, then this has consequences for programmers who are trying to construct or model such agents. Generally such a process of construction happens separately from the social situation that the agents are to inhabit – the programmer has a goal or specification in mind, tries to implement the agent to meet these and later the agents are situated so as to interact with others.

Whether this is *possible* to do depends on the extent to which the aspects of its intelligence are practically abstractable to a model which is analysable into two (essentially) unitary parts: the agent and its environment. If this can be done then one can indeed design the agent with this environment in mind. In this case the social environment is effectively modellable from the agent’s point of view.

If this sort of process is impractical (e.g. all the interactions in the social environment actually matter to the agent) this corresponds to a situation in which the agent is socially embedded [17]. Here the agent can not model its social environment as a whole and thus is forced to evolve heuristics based on the individuals it knows about in that environment. Some of these heuristics are listed below.

There are a number of possible responses to inhabiting such a social environment, including:

- Developing ways of structuring relationships to make them more reliable/predictable, e.g. contracts and friendship networks;
- Developing constraints on normal social behaviour via social norms and enforceable laws [5];
- Developing institutions and groupings that act to ‘filter out’ the complexity of the exterior social environment [2];

- To try and identify good sources of information and opinion and rely on these as a basis for decision making;
- To imitate those agents who many others imitate;
- To frequently sample the social environment via ‘gossip’;
- and, finally, to develop ones heuristics over time from within the relevant society and so avoid having to infer them later.

In practice many models of socially interacting agents take one (or a limited selection of) these heuristics as the *basis* for their agents. This is fine as long as one does not then make the false step of *defining* a social agent on the basis of one such heuristic. It is likely that intelligent and creative social agents that co-evolve within a society of other such agent that are individually recognisable will develop a many separate and different heuristics [16]. The heuristics are merely a *result* of being such an agent in such an environment. This leads us to believe that a bottom-up (or constructivist [15, 17, 20, 35, 37]) approach may be more profitable to a top-down *a priori* approach.

Challenges in SSI research

This section outlines a few research topics which we consider important to SSI research and which have in our view not yet gained as much attention as they deserve in the current research landscape. The list is not meant to be complete.

Culturally Situated Agents

The intelligent agents community which consists of people building software or hardware agents, or modelling societies of agents which show certain(social) intelligence, has so far not paid much attention to the issue that all technological products reflect the culture from which they originate. In the following we like to consider autonomous agents, following the definition given by Franklin and Grasser [19]: “An autonomous agent is a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future’. Currently, a paradigm shift from algorithms to interaction is acknowledged, see [38] which argues that recent technology is more than a continued development of more and more powerful rule-based Turing machines based on the closed-system metaphor. Instead, interactive systems, interaction machines which are inherently open systems are supposed to be the computational paradigm of the future. The shift of attention from algorithms to interaction also indicates a shift in beliefs, from the belief to discover and implement a universally intelligent, and general purpose machine towards an interactive machine, i.e. an agent which is not intelligent by itself by only behaves intelligently during interaction with other agents of the same or different kind, e.g. which interact with humans. Thus, the (social) context strongly matters, and in the case of interactions with humans such a socially situated agent which is used in different countries and communities also has to be a culturally situated agent [32] and a PRICAI98 workshop which addresses this topic[3]. We cannot expect that agents, both natural and artificial, behave identically in different social and cultural contexts. Thus, design and evaluation of agents could benefit from considering these issues.

---

[3] http://www.nttmsc.com.my/kido/pricai98cfp.html

Imitation and the ‘like-me’ test

A workshop at the latest Autonomous Agents AA’98 conference characterize simulation as follows: “Imitation is supposed to be among the least common and most complex forms of animal learning. It is found in highly socially living species which show, from a human observer point of view, ‘intelligent’ behaviour and signs for the evolution of traditions and culture. There is strong evidence for imitation in certain primates (humans and chimpanzees), cetaceans (whales and dolphins) and specific birds like parrots. Recently, imitation has begun to be studied in domains dealing with such non-natural agents as robots, as a tool for easing the programming of complex tasks or endowing groups of robots with the ability to share skills without the intervention of a programmer. Imitation plays an important role in the more general context of interaction and collaboration between agents and humans, e.g. between software agents and human users. Intelligent software agents need to get to know their users in order to assist them and do their work on behalf of humans. Imitation is therefore a means of establishing a ‘social relationship’ and learning about the actions of the user, in order include them into the agent’s own behavioural repertoire[4].

---

[4] `http://www.cis.udel.edu/~agents98/workshops/interaction.html`

Imitation is on the one hand considered an efficient mechanism of social learning, on the other hand experiments in developmental psychology suggest that infants use imitation to get to know persons, possibly applying a 'like-me' test ('persons are objects which I can imitate and which imitate me'), see discussion in [12]. Imitation (e.g. as social reinforcement techniques or programming by demonstration set-ups in robotics and machine learning) has been used primarily by focusing on the 'technological' dimension (e.g. imitation providing the context of learning sequences of action), and disregarded the social function of imitation. Additionally, the split between imitation research in natural sciences and the sciences of the artificial are difficult to bridge, we are far from a common research framework supporting an interdisciplinary approach toward simulation, cf. [31] for an attempt to provide a mathematical framework to facilitate analysis and evaluation of imitation research. With an embodied system inhabiting a non-trivial environment imitation addresses all major AI problems from perception-action coupling, body-schemata, recognition and matching of movements, reactive and cognitive aspects of imitation, the development of sociality, or the notion of 'self'), just to mention a few issues. Imitation involves at least two agents sharing a context, allowing one agent to learn from the other. The exchange of skills, knowledge, and experience between natural agents cannot be done by brain-to-brain communication in the way how computers can communicate via the internet, it its mediated via the body, the environment, the verbal or non-verbal expression or body language of the 'sender', which in return has to be interpreted and integrated in the 'listener's' own understanding and behavioural repertoire. And, as imitation games between babies and parents show, the metaphor of 'sender' and 'receiver' is deceptive, since the game emerges from the engagement of both agents in the interaction (see notions of situated activity and interactive emergence [22].

New Trends in Social Robotics: From sorting ants to soccer playing robots

In [8] Rodney Brooks gives a 'historical' overview on the shift of viewpoint from classical to behaviour-based robotics. Brooks was one of the strongest proponents of this shift of viewpoint or paradigm. Linked to the availability of simple and relatively inexpensive robots the new paradigm allowed to study robot group behaviour, instead of the classical approach which often focused on one monolithic system. A decade of research along the behaviour-oriented and artificial life approach has resulted in a number of impressive experiments demonstrating the emergence of group behaviour based on the interaction of simple robots performing simple behaviours, see [28] for an overview. For much of this work insect-like, in particular ant-like behaviour has been implicitly or explicitly served as a biological model. However, there are limits to the behavioural complexity one can achieve with this approach when trying to go beyond wall-following, flocking, herding, collective sorting, etc. In [8] again Rodney Brooks gives a strong push towards the direction of cognitive robotics. Rather than insects, mammals or even humans become popular models. Issues like memory and history, see [14] and [29] are among the currently investigated issues.

Recently a particular ‘application domain’ has gained a lot of attention in the autonomous agents community, namely the RoboCup[5] (see [25]). Teams of software and robotic agents join a competition and have to cope with real world constraints (e.g. noise) and limited resources (e.g. time constraints). Imitating human soccer playing is the target, and therefore cognitive aspects like individual roles, strategies, teamwork and cooperation (see [36]) have to be combined with the well-known low level behaviours like target following or obstacles avoidance. Thus, for those who cannot effort to buy or build a humanoid robot, the RoboCup challenge allows to tackle cognitive robotics on the team level! Additionally, modelling human soccer playing with autonomous robots (or software agents) opens the field of autonomous agents towards other field like computational organisation theory (see an attempt towards a symbiosis of both in [30].

References

1. Baron-Cohen, S., A. M. Leslie, U. Frith (1985) Does the autistic child have a ‘theory of mind’? *Cognition*, 21, 37-46
2. Bednarz, J. 1984. Complexity and Intersubjectivity. *Human Studies*, 7:55-70.
3. Beer, R. D. (1990). *Intelligence as Adaptive Behaviour*. Academic Press.
4. Billard, A., Hayes, G. and Dautenhahn, K. (1998). Experiments on human-robot communication with Robota, an imitative learning and communication doll robot. This volume.
5. Boman, M and Verhagen, H (1998). Social Intelligence as Norm Adaption. This volume.
6. Breazeal, C and Velasquez, J (1998). Toward teaching a robot ‘infant’ using emotive communication acts. This volume.
7. Brooks, R. (1991). Intelligence without Representation. *Artificial Intelligence*, 47:139-159.
8. Brooks, Rodney (1996) Behavior-based humanoid robotics, *Proc. 1996IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS96, 1-8.
9. Carley, K. and Newell, A. (1994). The Nature of the Social Agent. *Journal of Mathematical Sociology*, 19:221-262.
10. Carneiro, R. L. (1987). The Evolution of Complexity in Human Societies and its Mathematical Expression. *International Journal of Comparative Sociology*, 28:111-128.
11. Channon, A. D. and Damper, R. I. (1998). The Evolutionary Emergence of Socially Intelligent Agents. This volume.
12. Dautenhahn, K. (1995), Getting to know each other -- artificial social intelligence for autonomous robots, *Robotics and Autonomous Systems,* 16:333-356.
13. Dautenhahn, K. (1997). I Could Be You: The phenomenological dimension of social understanding. *Cybernetics and Systems*, 28:417-453.
14. Dautenhahn, K. and Chrystopher Nehaniv (1998) Artificial Life and Natural Stories, In: *International Symposium on Artificial Life and Robotics* (AROB III'98 - January 19-21, 1998, Beppu, Japan), vol 2,435-439
15. Drescher, G. L. (1991). *Made-up Minds – A Constructivist Approach to Artificial Intelligence.* Cambridge, MA: MIT Press.

---

[5] `http://www.robocup.org/`

16. Edmonds, B. (forthcoming). Modelling Socially Intelligent Agents. *Applied Artificial Intelligence*.
17. Edmonds, B. (1998). Capturing Social Embeddedness: a constructivist approach. CPM Report 98-34, MMU, UK. http://www.cpm.mmu.ac.uk/cpmrep34.html
18. Franklin, S. (1996). Coordination without Communication. University of Memphis, 1996.
19. Franklin, Stan and Art Graesser (1997) Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents, *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages*, published as Intelligent Agents III, pp 21-35, Springer-Verlag
20. Glasersfeld, E. von (1995). *Radical Constructivism: A Way of Knowing and Learning*. London: the Falmer Press.
21. Grassé P. P. (1959). La reconstruction du nid et les coordinations inter-individuelles chez Bellicositermes natalensis et Cubitermes sp. La theorie de la stigmergie: Essai d'interpretation des termites constructeurs. *Insect Societies*, 6:41-83.
22. Hendriks-Jansen, Horst (1996) *Catching ourselves in the act: situated activity, interactive emergence, evolution, and human thought*, MIT Press, Cambridge, Mass.
23. Hendriks-Jansen, Horst (1997) The Epistemology of Autism: Making a Case for an Embodied, Dynamic, and Historic Explanation, *Cybernetics and Systems*, 28:359-416
24. Jacobson, R. (1998). Positioning the Analysis: memetics as a methodological tool to close the ranks between social and traditional history. This volume.
25. Kitano, H. and M. Asada and Y. Kuniyoshi and I. Noda and E. Osawa (1997). RoboCup: The Robot World Cup Initiative, *Proc. of the First International Conference on Autonomous Agents*, Marina del Rey, CA, USA, February 5-8, ed.W. Lewis Johnson, 340-347
26. Kummer, H., Daston, L., Gigerenzer, G. and Silk, J. (1997). The social intelligence hypothesis. In Weingart et. al (eds.), *Human by Nature: between biology and the social sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 157-179.
27. Lawson, T. (1997). Situated Rationality. *Journal of Economic Methodology* 4:101-126.
28. Mataric, M. J. (1995). Issues and approaches in design of collective autonomous agents, *Robotics and Autonomous Systems*, 16:321-331
29. Michaud, François and Maja J Mataric, “Learning from History for Behavior-Based Mobile Robots in Non-stationary Conditions”, joint special issue on Learning in Autonomous Robots, *Machine Learning,* 31(1-3), 141-167,and Autonomous Robots, 5(3-4), Jul/Aug 1998, 335-354.
30. Moss, Scott and Kerstin Dautenhahn (1998): Hierarchical Organisation of Robots:a Social Simulation Study. *Proceedings 12th European Simulation Multiconference* ESM98, Manchester, United Kingdom June 16-19, 1998, 400-404.
31. Nehaniv, Chrystopher and Kerstin Dautenhahn (1998): Mapping between Dissimilar Bodies: Affordances and the Algebraic Foundations of Imitation. *Proceedings European Workshop on Learning Robots* 1998 (EWLR-7), Edinburgh,20 July 1998, editors: John Demiris and Andreas Birk.
32. O'Neill-Brown, Patricia (1997): Setting the Stage for the Culturally AdaptiveAgent. In: *Socially Intelligent Agents, Papers from the 1997 AAAI Fall Symposium* Technical Report FS-97-02. Menlo Park: The AAAI Press. 93-97.
33. Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Ballentine.
34. Prem, Erich (1998): The implications of embodiment for cognitive theories. A systems view, Proceedings of the 14th European Meeting on Cybernetics and Systems Research, EMCSR'98, 687-692
35. Riegler, A. (1992). Constructivist Artificial Life and Beyond. *Workshop on Autopoiesis and Perception*, Dublin City University, Aug. 1992.
36. Tambe, M. Implementing agent teams in dynamic multi-agent environments, Applied Artificial Intelligence 1998; volume 12

37. Vaario, J. (1994). Artificial Life as Constructivist AI. *Journal of SICE*.

38. Wegner, Peter (1997) The Paradigm Shift from Algorithms to Interaction, *CACM*, 40:80-91

39. Werner, E. Towards a Formal Theory of the Ontogeny of the Social Self.

40. Wimsatt, W. (1972). Complexity and Organisation, in Scavenger and Cohen (eds.), *Studies in the Philosophy of Sciences*. Dordrecht: Riddle, 67-86.

41. Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

Socially Intelligent Agents and The Primate Social Brain – Towards a Science of Social Minds

Kerstin Dautenhahn
Adaptive Systems Research Group
Faculty of Engineering & Information Sciences
University of Hertfordshire
Hatfield, Herts AL10 9AB, United Kingdom
K.Dautenhahn@herts.ac.uk
http://homepages.feis.herts.ac.uk/∼comqkd

Abstract

This article[1] puts research on socially intelligent agents (SIA) in the broader context of how humans (and other primates) perceive and interact with the social world. Phylogenetic (evolutionary) and ontogenetic (developmental) issues are discussed with respect to the social origin of primate and human intelligence and human culture. Implications for designing artifacts and for the evolvability of human societies are outlined. A theory of empathy is presented that is based on current research on the primate social brain. Research projects that investigate some of these issues are reviewed. I argue that Socially Intelligent Agents (SIA) research, although strongly linked to software and robotic engineering, goes beyond a software engineering paradigm: it can potentially serve as a paradigm for a science of social minds. A systematic and experimental investigation of human social minds and the way humans perceive the social world can result in truly social artifacts, that are integrated in human society.

"*Once there lived a robot called Rob, he was made by a mad professor called Brain-Box on Jan the 3rd 2,000,000 in Germany. Everybody thought it was brilliant and on Jan 30th the same year it ran away and it ran away to England and terrorised England people looked to see who or what had done it. The robot had killed someone and taken her brain they got the police out and they looked into it the police didn't know who or what had done it. One day someone called Tod saw Rob and called the police to come and get Rob but when the police cam Rob had gone. Tod got a fine for supposedly lying to the police. The next year the police had still not found Rob. Sometime in the last year Rob had fallen off something and he was found in pieces in a junk yard. Some people phoned the police the police came out and he had gone. The village dug a massive pit to try and catch Rob. Now Rob was as high as six houses he was eating a lot of junk metal. One day rob was walking along and he fell in the pit, the people found him in the pit and he was killed.*" (Christopher, 9 years old, (BD99)).

[1]This is a slightly updated version of a paper published as *Proc. AAAI Fall Symposium "Socially Intelligent Agents - The Human in the Loop"*, AAAI Press, Technical Report FS-00-04, pp. 35-51, November 2000.

Introduction: Socially Intelligent Agents - The Human in the Loop

As Reeves and Nass have shown (RN96) humans tend to treat computers (and media in general) as people. I believe that this 'media equation' (media equals real life) is particularly relevant for socially intelligent agents (SIA's) research with the 'human in the loop', namely studying the relationship between socially intelligent agents and humans as designers, users, observers, assistants, collaborators, competitors, customers, or friends. In order to acknowledge the 'human in the loop' I suggested in (Dau98) a list of design guidelines for SIA technology, identifying the following roles of humans and suggesting that a balanced design of socially intelligent agents need to address these roles: Humans are embodied agents, humans are active agents, humans are individuals, humans are social beings, humans are storytellers, humans are animators, humans are autobiographic agents, humans are observers. Not unsurprisingly, SIA research is more than other agent research strongly inspired and motivated by findings from outside software engineering and computer science, in particular by the humanities, social sciences, and natural sciences. As such, SIA research is different from the field of agent-based computing research that views agents primarily as a software engineering paradigm (cf. (JSW98), (WJK00)). In contrast, SIA research benefits from viewing agents in the larger picture of autonomous agents as defined by Franklin and Graesser ((FG97)): "An autonomous agent is a system situated within and a part of the environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to affect what it senses in the future". This definition is very attractive since it applies easily both to natural and artificial systems. Franklin and Graesser (FG97) also propose a natural kinds taxonomy of agents, distinguishing between biological, robotic and computational agents as instances of the class of autonomous agents. In a different paper Stan Franklin argues for the study of autonomous agents as embodied artificial intelligence (Fra97)[2]. I believe that this was an important step, namely viewing (autonomous) agents as vehicles and subjects for the investigation of artificial (and human) intelligence, and in this paper I like to argue that socially intelligent agents can similarly serve as tools and vehicles for the study of artificial (and human) social minds.

Socially Intelligent Agents research is concerned with agents and human-agent interactions whereby the agents show aspects of human-style social intelligence.

> Socially intelligent agents are biological or artificial agents that show elements of (human-style) social intelligence. The term *artificial social intelligence* refers then to an instantiation of human-style social intelligence in artificial agents. (Dau98)

Social intelligence can be natural (humans) or artificial (computational or robotic agents), but within the context of human-style social interaction and behaviour. Please note that social intelligence in this sense does not make claims on how intelligent the agent needs to be: often simple strategies can be socially very effective!

Although SIA research is primarily interested in human-style behaviour and interactions, human social intelligence has a history, both evolutionary as well as developmental. Below I therefore discuss a few findings from primatology and developmental psychology and their implications for SIA research.

The social, ethical, cultural, as well as cognitive implications of SIA technology are important issues to consider. Even on the level of the individual human, interactions with SIA's can influence a human beings attitudes, behaviour and minds, and in this way empower as well as manipulate humans (see discussion in (ND00). In (Fog99) B.J. Fogg discusses computers as *persuasive technologies*. In contrast to other non-persuasive technologies, "persuasive computing technology is a computing system, device, or application intentionally designed to change a person's attitudes or behaviour in a predetermined way." Furthermore, Fogg calls the study of planned persuasive effects of computer technologies *captology*. Figure 1 shows the functional triad of computer persuasion. Following Fogg's terminology SIA's might fall under the category of 'social actors', where agents can adopt animate characteristics, play animate roles, and follow social dynamics for the purpose of creating relationships with humans and invoke social responses. In this sense, SIA's are persuasive technologies and therefore issues of design, credibility (TF99), and ethics of persuasive technology (BN99) also apply to SIA's technology, in particular to the new generation of highly interactive 'social' software and robotic agents, many of them described in this volume and elsewhere.[3].

The Life-Like Agents Hypothesis

As discussed in the previous section, SIA's are often designed to 'imitate' life. Based on what I called previously the 'Life- Like Agents Hypothesis' this approach can be characterised as follows (Dau99c):

> "Artificial social agents (robotic or software) which are supposed to interact with humans are

[2]Embodiment is here considered as 'embodied in the situated sense of being autonomous agents structurally coupled with their environment'. In this sense software agents can be as embodied as robotic and biological agents. Tom Quick, Kerstin Dautenhahn and Chrystopher Nehaniv developed a definition of embodiment based on structural coupling and mutual perturbation between an agent and its environment, a definition that applies to different kinds of systems, including autonomous agents, and which allows to measure different degrees of embodiment quantitatively ((QDNR99b), (QDNR99a), (QDNR99c)), see a more recent discussion in K. Dautenhahn, B. Ogden, T. Quick: From Embodied to Socially Embedded Agents - Implications for Interaction-Aware Robots, to appear in *Cognitive Systems Research*, special issue on Situated and Embodied Cognition, guest-editor: Tom Ziemke, Elsevier.

[3]Examples of collections of articles on SIA research are: K.Dautenhahn, C. Numaoka (guest editors): Socially Intelligent Agents, Special Issues of *Applied Artificial Intelligence*, Vol. 12 (7-8), 1998, and Vol. 13(3), 1999, K. Dautenhahn, ed. (2000): *Human Cognition and Social Agent Technology*, John Benjamins Publishing Company, B. Edmonds and K. Dautenhahn (guest editors): Social Intelligence, special issue of *Computational and Mathematical Organisation Theory*, Vol. 5(3), 1999, K. Dautenhahn (guest editor): Simulation Models of Social Agents, special issue of *Adaptive Behavior*, Vol. 7(3-4), 1999, Bruce Edmonds and Kerstin Dautenhahn (guest editors): Starting from Society - the application of social analogies to computational systems, special issue of *The Journal of Artificial Societies and Social Simulation (JASSS)*, 2001. Kerstin Dautenhahn (guest editor): Socially Intelligent Agents – The Human in the Loop, special issue of *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, Vol. 31(5), 2001; Lola Cañamero and Paolo Petta (guest editors), Grounding emotions in adaptive systems, special issue of *Cybernetics and Systems*, Vol. 32(5) and Vol. 32(6), 2001. K. Dautenhahn, A. Bond, L. Cañamero, B. Edmonds, eds. (2002): *Socially Intelligent Agents - Creating Relationships with Computers and Robots*, Kluwer Academic Publishers. K. Dautenhahn, C. L. Nehaniv, eds (2002): *Imitation in Animals and Artifacts*, MIT Press. C. L. Nehaniv, K. Dautenhahn (guest editors): Imitation in Natural and Artificial Systems, special issue of *Cybernetics and Systems*, Vol. 32(1-2) 2001.

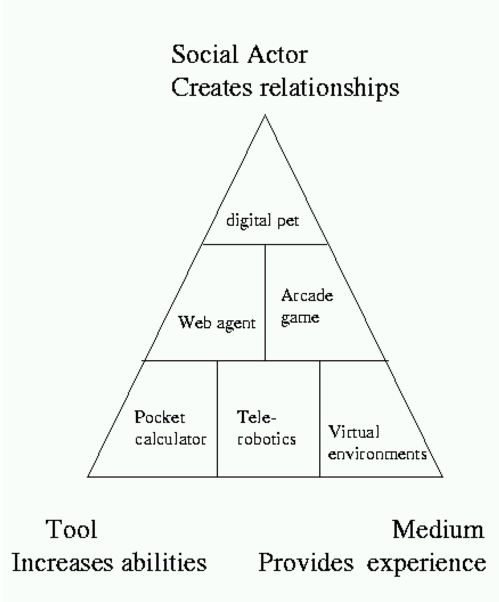

Figure 1: The functional triad of computer persuasion, redrawn from (Fog99).

most successfully designed by imitating life, i.e. making the agents mimic as closely as possible animals, in particular humans. This comprises both 'shallow' approaches focussing on the presentation and believability of the agents, as well as 'deep' architectures which attempt to model faithfully animal cognition and intelligence. Such life-like agents are desirable since 1) The agents are supposed to act on behalf of or in collaboration with humans; they adopt roles and fulfill tasks normally done by humans, thus they require human forms of (social) intelligence, 2) Users prefer to interact ideally with other humans and less ideally with human-like agents. Thus, life-like agents can naturally be integrated in human work and entertainment environment, e.g. as assistants or pets, and 3) Life-like agents can serve as models for the scientific investigation of animal behaviour and animal minds".

Argument (3) is certainly valid and need not be discussed here. However, arguments (1) and (2) are not as straightforward as they seem. Designing life-like agents that closely mimic human appearance or behaviour can unnecessarily restrict and narrow the apparent and actual functionality of an agent. Similarly, imagine mobile phones were designed so that they had the shape of old-fashioned dial-operated telephones. It then could be disturbing or at least puzzling for people to find out that the mobile phone might have more functionalities (e.g. sending and receiving email, browsing the Web etc.) than the original model. Thus, 'new' designs not imitating any other previously existing object might better suit a piece of technology that is combining functionalities in a novel way or has new functionalities. A social interface agent (e.g. in an e-commerce context) presented with humanoid appearance and behaviour might have the advantage of evoking an initial feeling of 'familiarity' in a human customer, but 1) human customer's are then likely to expect the agent to show other human characteristics and functionalities, human knowledge, personality and other characteristics of humans in general (including that it understands jokes and possesses common sense knowledge), and sales agents in particular, and 2) new or different functionalities that the real agent does not possess need to be integrated in a plausible way in the agent's behaviour, without breaking the suspense of disbelief (Mat97) (see (ND00) for further discussion of these issues).

Attitudes towards Socially Intelligent Agents: Anthropomorphism and Behaviour Reading

According to the *Social Intelligence Hypothesis*, discussed in more detail below, the evolution of primate intelligence is linked with an increase of the complexity of primate social life ((BW97), (WB88)). The argument suggests that during the evolution of human intelligence a transfer took place from social to non-social intelligence so that hominid primates could transfer their expertise from the social to the non-social domain (see review in (Gig97)). An interesting aspect of this kind of transfer is given by Mithen (Mit96), who explains the evolution of anthropomorphic thinking with an accessibility between the domains of social intelligence and natural history intelligence so that "people could be thought of as animals, and animals could be thought of as people", (Mithen 1996, p. 224).

The attribution of human motivation, characteristics, or behaviour to inanimate objects, animals or natural phenomena is usually called anthropomorphism (see The American Heritage©Dictionary of the English Language). Anthropomorphism is often dismissed as a curiosity or unscientific phenomenon and only relatively few scientific work outside philosophy have experimentally addressed the issues of how and why people tend to adopt an *intentional stance* (Den71), (Den87), namely treating non-humam objects and animals as intentional objects (what seems to be based on the human mindreading or social competence system, see discussion below). It is often suggested that physical likeness, familiarity, phylogeny and/or cultural stereotypes are important factors. Well known is the study by Eddy et al. (EGP93) who investigated peoples tendency to anthropomorphise animals (see summary in (Wat97)). The study identified two primary mechanisms why people attribute similar experiences or cognitive abilities to animals, based on 1) the degree of physical similarity, and 2) the degree of an existing attachment bond (familiarity). Dogs and cats are more familiar to most people than frogs, primates are physically (and behaviourally) similar to humans.

This study seems to support the above mentioned *Life-Like Agents Hypothesis*, namely that humanoid agents that look like humans should be more believable and successful as social interaction partners for humans than non-humanoid agents (assuming that humans mostly enjoy interacting with other humans). However, other evidence suggests that not physical similarity, but *behaviour in context* matters. Mitchell and Hamm (MH97) provided undergraduate students with narratives depicting different mammalian agents (including humans) showing behaviour that suggested jealousy or deception. The students were then asked to answer questions on particular psychological characterisations of the agents. The narratives varied according to species, context in which an agent's behaviour occurred, and the degree of emphasis that the narrative was about a particular species of animal (or human). The behaviour was constant in all narratives. Mitchell and Hamm found that variations in context influenced the psychological characterisations, but variations in species and emphasis did not, i.e. the psychological characterisations of all species were almost always similar: "Nonscientists (and some scientists as well) apparently use a mammal's behavior-in-context (whether human or not) as evidence of its psychological nature, regardless of the mammal's physical similarity, familiarity, or phylogenetic closeness to humans, or the mammal's cultural stereotype; psychological terms are not used specifically for humans, but rather are depictive of behaviour-in-context". Interestingly, the notion that behaviour matters more than appearance in ascribing intentionality is supported by an experimental study published in 1944 (HS44) that convincingly demonstrates the effects of the 'intentional stance'. Here, human subjects created elaborate narratives about intentional agents when asked to describe movements of moving geometric shapes shown in a silent film. Other studies along research done by Mitchell and Hamm and Heider and Simmel could confirm whether this also applies to non-mammalian animals. A particular challenge would be to include computational and robotic agents in such studies. I suggest that behaviour-reading might apply also to inanimate objects such as robots. Every robotics researcher who has ever given a demonstration of autonomous mobile robots to a general audience can confirm how readily humans view robots as people, cf. (Bra84), (BD99). The importance of behaviour expression in agent building has been recognised e.g. by Phoebe Sengers (Sen98), (Sen00). Her argument is that 'doing the right thing' (the classical approach of AI approaches to agent control) needs to be complemented by paying attention to 'doing the thing right', in particular creating believable transitions between agent behaviours.

Attitudes Towards Agents: A Case Study with Robots

In (BD99) Kate Bumby and Kerstin Dautenhahn investigated children's attitudes towards robots, a brief summary is given here. We were interested to find out how children interact and describe robots. Thirty eight children (ages seven to eleven, 21 males, 17 females, BC1 socioeconomic category) were studied at St. Margarets Junior School in Durham, UK. A number of working hypotheses were addressed with respect to how the children portrayed robots. In three studies the children were asked a) to draw a picture of a robot, b) to write a story about the robot they had drawn. These studies were observational. The third study had the format of an informal, guided and filmed interview while the children were in a group interacting with two mobile robots (see figure 2) that were running in an environment with a light source. The robots were simple behaviour-based vehicles (Bra84).

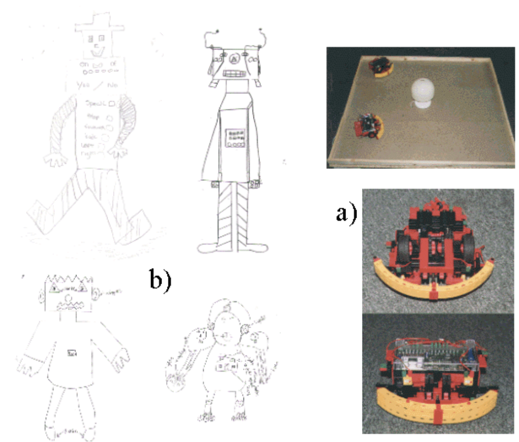

Figure 2: a) The experimental set up and the two autonomous, mobile *fischertechnik* robots, b) Drawings of 8-year and 9-year olds, (BD99).

Results of study a (pictures) show e.g. that the children tend to give the robot humanoid faces. Figure 2 shows examples of a variety of drawings by 8/9-year olds, portraying robots. In study b (stories), one result was that the children tend to put the robots in familiar settings, doing familiar tasks. The robots were significantly often put in a social context. Study c (interview) showed a clear tendency to anthropomorphise the robots, e.g. "I don't think it likes the light.". The children also often talked to the robots as if they were animals or small children. Other findings of this case study, e.g. with respect to attribution of gender or violence is reported in more detail in (BD99). This single case study cannot answer the question of how children in that age range in general think about robots, but the results give some indication that confirm findings along the lines of studies with computers (RN96).

Societies of Social Animals[4]

Swarm Intelligence: Social Insects Don't have Friends

The term 'societies' is generally applied both to human and other animal societies, including social insects. Social insects (e.g. termites, bees, ants) are very well studied and two important theoretical concepts are used to understand coordination in social insect societies, namely *self- organisation* and *stigmergy*. Recently, models of *swarm intelligence* and their applications to problems like combinatorial optimisation and routing in communications networks have been studied extensively (BDT99). The concept *stigmergy* describes a class of mechanisms mediating animal-animal interactions, based on the description of insect behaviour as stimulus-response (S-R) sequences (even for solitary species). Stigmergy is based on indirect communication, communication via the environment, and an example of *collective behaviour*.

Primate Intelligence: Getting to Know Each Other

In primate societies, and different from members of social insect societies, an individual is not only socially situated (being part of and surrounded by a social environment) but also socially embedded (ED98) which means that the agent needs to pay attention to other agents and their interactions individually. Particularly human primates are specialised in predicting, manipulating and dealing with highly complex social dynamics (involving direct relationships as well as third-party relationships); they possess language as an effective means of preserving group coherence, 'social grooming' (Dun93), (BD97) and communicate about themselves and others in terms of stories (Dau99b). Humans are not only dealing with very complex relationships but seem to have mental 'models' of themselves, others and the social world (cf. (Whi91), (BC95)). Humans, different from social insects live in *individualised societies* (as do some other species of birds and mammals). An increasingly complex social field and an increasing need to effectively communicate with each other were likely to have been among the important constraints in the evolution of human minds.

Minds are certainly attributed to members of *Homo sapiens* (and as some evidence suggests several other hominid species might have existed with 'minds'), but other candidates exist among mammals (e.g. nonhuman apes, dolphins, elephants) and birds (e.g. parrots and members of the crow family). Interestingly, species which we describe as possessing a 'mind' are all highly social. Even the 'solitary' life style of *Pongo pygmaeus* or orangutans, (who nevertheless seem to be highly social in their ability to recognise and interact with each other) is rather a secondary adaptation to a particular environment which demands a spatially 'distributed' social organisation. The *Social Intelligence Hypothesis* suggests that primate intelligence primarily evolved in adaptation to social complexity, i.e. in order to interpret, predict and manipulate conspecifics (see overview in (BW97), (WB88)). Thus, there are two important aspects to human sociality: it served as an evolutionary constraint which led to an increase of brain size in primates, this in return led to an increased capacity to further develop social complexity.

Although it is still unknown why hominids needed or chose to live in social groups, this *feedback principle* soon led to the development of highly sophisticated levels of organisation and control in human societies. In (Rus93) four levels of primate social organisation are discussed which might serve as models for the evolution of primate societies: a) the 'shrew'-type pre- primates: solitary, many offspring, insectivores, e.g. *Purgatorius*, a 70- million-year-old fossil, b) the 'mouse-lemur'-type primates: bush-living, nocturnal, strong mother-daughter bonding (stable matrilines), social learning (offspring learns from mother), solitary males and social groups of mothers and daughters, e.g. the 50-million-year-old fossil *Shoshonius cooperi*, c) the '*Lemur catta*'-type diurnal lemurs: appearing about 54 million years ago, social groups (troops), dominant females, submissive males, stable matrilines, occasionally consort bonds between single male and female, e.g. Adapidae, d) the 'chimpanzee'-type lemur-ape: appearing about 24 million years ago, groups of dominant males and submissive females, stable families of mothers and their offspring, male power coalitions, e.g. *Dryopithecus*. The social organisation of recent species of apes shows variations of this pat-

[4] This section is based on (Dau00b).

tern: of harem-structures (gorilla), solitary lifestyle (orangutan). Such stages of social organisation can be related to behavioural as well as cognitive capacities of primates.

The terms 'theory of mind' and *mindreading* are usually used in order to discuss whether an animal is able to reflect on its own mental states (e.g. desires, intentions and beliefs) and those of others. Researchers have studied whether humans might have particularly specialised in a theory of mind (PW78), (PP95a). However, as Richard Byrne pointed out (Byr97), the *Social Intelligence Hypothesis* might account for the evolution of primate intelligence, but not for the specific human kind of intelligence. Here, narrative psychology and studies on the development of autobiographic memory and a 'self' might offer an explanation: evidence suggests that 'stories' are the most efficient and natural human way to communicate, in particular to communicate about others (Bru91). Dennett (Den89) even regards the 'self' as a 'centre of narrative gravity'. Narrativity, the capacity to communicate in terms of stories is therefore regarded an efficient means to communicate social matters, and the origin of narratives might therefore have been a crucial milestone in the evolution of primate social intelligence (RM95). The *Narrative Intelligence Hypothesis* (NIH) (Dau99b) proposes that the evolutionary origin of communicating in stories was correlated with increasing social dynamics among our human ancestors (see figure 3), in particular the necessity to communicate about *third-party relationships* (which in humans reaches the highest degree of sophistication among all apes, cf. gossip)[5].

A number of current research is devoted to building narrative software, virtual or physical environments (e.g. (GC97), (MMP99), (BC00), (BID+99), (MAD+00), (BBA+00)). Supporting human narrative intelligence is expected to impact human minds and our notions of sociality and what we call our 'selves'. In parallel, investigations into autonomous story-telling agents can result in agents (robotic or software) with genuine narrative minds, being able to tell us interesting stories, listen to and understand our stories, and make us laugh. A first attempt to a bottom-up approach to narrative intelligence for autonomous agents is described in (CD00), (DC00). The kind of stories these agents might tell us will be shaped by the social field and the cultural environment of human societies

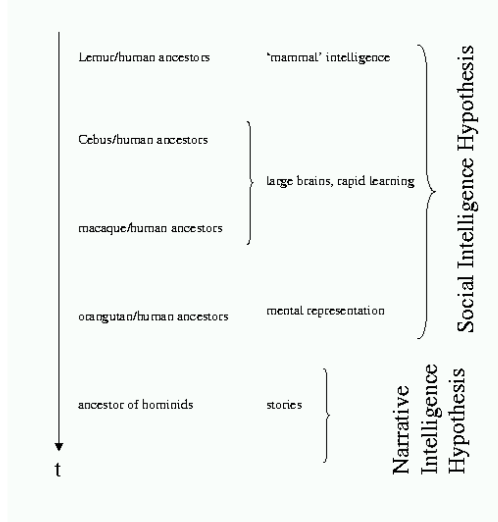

Figure 3: Evolution of the Social Mind, modified from (Byr97), see text for explanation.

in which these agents 'grow up'. Thus, it's up to us whether the stories of the future will be nightmares, fairy-tales, comedies or adventures.

> "*Once they were a robot, that lived in the country. He came to this cliff, he thought it was too steep for him to go down so he went down the steps. He went onto the beach oh yeah the robot's name is Shaped and Shaped went to play in the sea and after a few minutes he fell to the ground and got washed into shore because the water has gone into his body.*" (Becky - 8 years old, (BD99)).

Primate Culture: We are not alone

The terms *anonymous* and *individualised* societies are used in biology in order to describe two different types of social organisation. Social insects are the most prominent example of anonymous (eusocial) societies where group members do not recognise each other as individuals but rather as group members[6]. We do not observe bees or termites searching for missing members of their colony. Although individuals adopt specific roles in a colony they do not show individuality or 'personality'.

The situation is quite different in individualised societies which primate societies belong among. Here we find complex recognition mechanisms of kin and group members. This gives rise to complex kinds of social interaction and the development of various forms of social relationships and networks. On the

[5] For a more recent discussion of the NIH see K. Dautenhahn: The Origins of Narrative - In Search for the Transactional Format of Narratives in Humans and Other Social Animals, accepted for publication in *International Journal of Cognition and Technology: Co-existence, Convergence, Co-evolution (IJCT)*, John Benjamins Publishing Company

[6] Note that African naked mole-rats, mammals, show a eusocial organisation similar to social insects (SJA91). Thus, the eusocial form of organisation has evolved independently in different taxa of animals.

behavioural level long-lasting social bonding, attachment, alliances, dynamic (not genetically determined) hierarchies, social learning, development of traditions etc. are visible signs of individualised societies. In humans the evolution of language, culture and an elaborate cognitive system of mindreading and empathy are characteristics of human social intelligence in individualised societies (Dau97). As a consequence of the latter, humans are not only paying attention to other agents and their interactions individually (interactions between distinct personalities), but they use their mental capacities to reason about other agents and social interactions. It is at present unclear to what extent the social intelligence of members of other animal species, in particular very social species like elephants, Grey parrots, non-human apes and cetaceans, is similar to or different from our own. Similarly, the issues of cultural and 'memetic' evolution is highly controversial. The concept of memes, first introduced by Dawkins (Daw76) comprises ideas, fashions, skills and other components of *human* culture. Human culture and the memetic transmission of knowledge, ideas and skills is often regarded unique to human societies. According to Donald's discussion of the evolution of culture and cognition (Don93) modern humans have three systems of memory organisation (mimetic skill, language and external symbols) not available to our primate relatives, and these 'inventive capacities' result in languages, gestures, social rituals, images etc. According to Tomasello et al. (TKR93) cultural learning is a uniquely human form of social learning. Cultural learning requires three social-cognitive processes which emerge in human ontogeny: imitative learning, instructed learning (teaching) and collaborative learning. Similarly, Blackmore (Bla99) argues that only sophisticated forms of imitation which are characteristic of humans but not non-human primates, were a necessary prerequisite for memetic replication which leads to human culture.

Others argue that culture as such is unlikely to be a feature unique to human societies and that the acquisition of novel behaviours in 'proto-cultures' can be observed in animals. To give an example: traditions have been observed among troops of Japanese macaque monkeys (Huf96): Japanese macaques or *Macaca fuscata* show several examples of the acquisition of innovative cultural behaviours, e.g. sweet potato washing and wheat-washing was invented in 1953 by a young female and subsequently spreading to older kin, siblings, and playmates, eventually to other members of the troop. Other observed cultural behaviours are fish eating (as many newly acquired food sources initially spreading from peripheral males to adult females, then from older to younger individuals), and stone handling or stone play (initially spread only laterally among individuals of the same age). Subsequently all these behaviours were passed down from older to younger individuals in successive generations (*tradition phase*). These examples clearly show the influence of social networks on the *transmission phase* of novel behaviour: the nature of the behaviour and social networks determine how the behaviours are initially transmitted, depending on who is likely to be together in a certain context and therefore is exposed to the novel behaviour. Innovative behaviours of the kind described here have been independently observed at different sites. Various factors have been discussed which influence cultural transmission: environmental factors, gender, and age, and other social and biological life history variables. For example, unlike potato or wheat washing, stone handling declines when individuals mature.

With respect to cultural transmission in non-human apes, recent evaluations of long-term field studies of chimpanzees or *Pan troglodytes* give compelling evidence for cultural behavioural variants (traditions) in different chimpanzee communities, data which cannot be explained by ecological differences of the habitats, and comprising dozens of different behaviours including tool usage, grooming and courtship behaviours (WGM$^+$99). Possibly the kind of *mechanisms* that are necessary for and support culture (e.g. cognitive mechanisms, language, imitation) might be different in different animal species. As Frans del Waal concludes (dW99): "The 'culture' label befits any species, such as the chimpanzee, in which one community can readily be distinguished from another by its unique suite of behavioural characteristics. Biologically speaking, humans have never been alone - now the same can be said of culture.".

The striking similarity of cultural transmission of novel behaviour exhibited by Japanese macaque monkeys and chimpanzees and what we call human culture questions the uniqueness of human societies. Note that this behaviour is observed in monkeys, which do not show complex forms of social learning like imitation, and do not seem to possess higher-level 'cognitive' capacities necessary for complex social forms of 'primate politics' shown by non-human apes and humans. However, many non-human primates are very good social learners (widely using non-imitative forms of social learning, e.g. stimulus enhancement or social facilitation). Reader and Laland (RL99) therefore argue that the meme concept can and should also be applied to cultural transmission among non-human animals. Animal societies can appear in various forms. Human societies, human culture and human minds reflect in many ways their evolutionary origin in animal societies, animal culture and animal minds. Considering human culture in an evolutionary context, linking it to precursors in non-human primate societies might help a better understanding of human culture.

Implications for Evolvability of Human Societies

According to Kirschner and Gerhart evolvability can be defined as "the capacity to generate heritable phenotypic variation" (KG98). As outlined above, in primate non-human societies we already observe precursors of human culture (e.g. social learning, traditions). For reasons still under dispute our human ancestors were required to deal with increasingly complex social dynamics. Mental capacities evolved which allowed the evolution of increasingly complex mechanisms of social control, which in return increased the complexity of primate societies.

Based on what the previous sections discussed about primate societies and culture, the following requirements for mental capacities and social skills which facilitated the evolution of *primate culture* are proposed:

- Individualised societies: The capacity to identify and recognise individual group members.
- Social Networks: the capacity to establish, maintain, remember and utilise social networks. Three basic elements[7] in the primate social field are:
  - Forming direct (one-to-one) relationships with group members
  - Identifying third-party relationships (relationships among other group members)
  - Recognition of conspecifics as members in a group hierarchy/social structure (e.g. structures of kinship, allies, dominance hierarchies, etc.)
- Efficient mechanisms of social bonding, either via physical grooming (in on-human primate societies) or via language and communication in narratives as efficient ways of 'social grooming', important for maintaining the coherence of social groups at different levels of social organisation.
- Social learning: the capacity to use others as 'social tools' (as explained in (Dau95)), via social learning mechanisms with varying degrees of what the animals learn from each other (cf. social facilitation versus imitation)

Thus, culture and other characteristics of human societies cannot be separated from specific environmental (including social) constraints and mental capacities which evolved as adaptations for dealing with such constraints. Specific adaptations then turned out to be prerequisites in the evolution of more sophisticated forms of primate societies and culture. Although new forms of media seem to substantially expand the social life of humans, even today the same mental capacities which were involved in the evolution of the human social animal now pose cognitive limits on the complexity and number of social encounters. Our primate social brain has a limit on the number of individuals who we can maintain direct social relationships with, namely relationships based on direct social knowledge (around 150), correlated with the relative size of the human neocortex (Dun93), (BD97). This figure can be identified consistently in various ancient and present human cultures. This number is significantly larger, namely more than double that observed in any population of non-human primates. Unless drastic (technological) enhancements of human cognitive capacities are invented, this number could only be exceeded by inventing new, more efficient ways of "social grooming" (exceeding the communicative capacities of language). Another interesting issue discussed by Dunbar (Dun93), (BD97) is that language is 2.8 times more efficient as a mechanism of social bonding in comparison to physical grooming. The suggestion is therefore that human conversational group sizes should be limited to about 3.8 (which means one speaker and 2-3 listeners). Data on small group sizes confirm this hypothesis.

I showed above that biological evolution led to two distinctively different forms of social organisation in animal societies (anonymous and individualised societies). It appears that individualised societies were a necessary (but not sufficient) prerequisite for the evolution of culture, providing a social environment which supported the evolution of complex forms of social learning (in particular imitation). The capacity for phenotypic, cultural evolution seems correlated with particular mental capacities and social skills (see list above) which facilitated the evolution of complex forms of primate societies and primate culture. Primate social behaviour is well studied, we know less about the social life and mental capacities of non-primate species (crows, parrots, cetaceans, elephants, and others). However, when searching for animal culture, highly social animals in individualised societies are good candidates. Ants don't imitate, they don't learn from each other, primates do. Memes, as the replicators of culture, seem to require 'a social host', and memes are transmitted along social networks and depending on interactions its 'host' is engaged in. These seem to be the natural constraints under which culture is able to evolve in primate societies. The 'magic

[7] This is not supposed to be an exhaustive list.

numbers' 150 and 3.8 indicate strong limitations and constraints for the future development of human societies. Systematic investigations that take these cognitive constraints into consideration could provide a basis for social agent technology that meets the cognitive demands of human primates.

Social Robots in Rehabilitation: the Case of Autism[8]

Autism

Although we use the term autism throughout this paper it is more appropriate to use the term autistic spectrum disorders (ASD) which acknowledges the fact that autism occurs in differing degrees and in a variety of forms. The National Autistic Society (NAS00) lists the following triad of impairments: 1. Social interaction (difficulty with social relationships, for example appearing aloof and indifferent to other people, inappropriate social interactions, inability to relate to others in a meaningful way, impaired capacity to understand other's feelings or mental states). 2. Social communication (difficulty with verbal and non-verbal communication, for example not really understanding the meaning of gestures, facial expressions or tone of voice). 3. Imagination (difficulty in the development of play and imagination, for example having a limited range of imaginative activities, possibly copied and pursued rigidly and repetitively).

In addition to this triad, repetitive behaviour patterns and a resistance to change in routine can generally be observed, associated with a significantly reduced repertoire of activities and interests, stereotypical behaviour, and a tendency of fixation to stable environments. Depending on what is included in 'autism', rates of occurrence are given which range between 5-15 in 10000. Instead of a physical handicap which prevents people from physically interacting with the environment, people with autism have great difficulty in making sense of the world, in particular the social world. Autism can but need not be accompanied by learning disabilities. At the higher functioning end of the autistic spectrum we find people with Asperger Syndrome. Some of them manage to live independently as adults and to succeed in their profession, but only by learning and applying explicit rules in order to overcome the 'social barrier' (Gra95), (GS96), (Sch97). Instead of picking up and interpreting social cues 'naturally' they can learn and memorise rules about what kind of behaviour is socially appropriate during interaction with non-autistic people. Autism is not, as has long been assumed in public, a voluntary decision to retract from the world: people with autism do not have the choice to live socially or not, the decision has been made for them. Two different viewpoints exist on how to connect the autistic with the non- autistic world: either efforts are undertaken to teach people with autism the skills they need to survive in the world of 'normal' people, or it is suggested that they might be happier living separately in a world specifically designed for them. From all what we know about the way individuals with autism feel (see books written by Temple Grandin and others), they are painfully aware of their 'being different' from other people, and express the wish to be part of the 'world outside'. Accepting the differences, empowering people with autism, and linking their world with the world that non-autistic people are living in poses many challenges. In order to understand people with autism we have to understand better the causes of autism, and can find ways to empower them, including computer and robotic technology, so that they have the choice of whether and to what extent they want to connect to the world of non-autistic people.

[8]This section is based on (Dau00a).

Brief project Description and Related Work

The AURORA project develops an autonomous, mobile robot as a therapeutic tool for children with autism (Dau99c), (WD99), (DW00). Conceptually, this approach is strongly related to Seymour Papert's constructionist approach towards learning (Pap80). Such an approach focuses on active exploration of the environment, namely improvisational, self-directed, 'playful' activities in appropriate learning environments ('contexts') which can be used as 'personal media'. In the mid- 1960ies Papert and his colleagues at the MIT AI LAB developed the programming language LOGO which has been widely used in teaching children. A remote controlled device (a 'turtle' robot) was developed which is moving according to a set of LOGO instructions, cf. the LEGO/LOGO Artificial Life Toolkit for children (Res89). In 1976 Sylvia Weir and Ricky Emanuel (WE76) published research which used such a LOGO learning environment to catalyse communication in an autistic child. They report on their experience with a seven-year-old autistic boy and the positive effects of his explorations in controlling a LOGO turtle on his behaviour. A more recent approach using more interactive rather than remote-controlled technology for rehabilitation of autistic children is taken in the Affective Social Quotient (ASQ) project, (Blo99). Here, embedded technology is used to support autistic children in learning about social-emotional cues. Short 'emotionally charged' video clips are used together with a set of physical stuffed 'dolls' (embodying one emotional expression) through which the child can interact with the movies. By touching the doll the child can match a doll with a video clip. A child can explore emotional situations by picking up dolls with certain emotions, or the system can prompt the child to pick up dolls that go with certain clips. A therapist is able to control and monitor the interactions. The system shows that human-intensive, repetitive aspects of existing behavioural therapy techniques can potentially be automated.

In recent years the AURORA project described below and work by Francois Michaud (Michaud, this volume; (MCLL00), (MLL+00)) who develops interesting interactive robotic designs, is taking up this line of work. Since end of 1998 the project AURORA (AUtonomous RObotic platform as a Remedial tool for children with Autism) investigates how an autonomous mobile robot can be developed into a remedial tool in order to encourage children to become engaged in a variety of different interactions that possess features which are important elements of human social behaviour (eye-contact, joint-attention, approach, avoidance, following, imitation games etc.). The children who are interacting with the robot are between 8-12 years of age, including children who are non-verbal, i.e. they cannot use language or usually do not use language. In the rehabilitation of children with autism therapeutic issues (e.g. eye contact, joint attention, turn taking, reading mental states and emotions) are usually addressed in constrained teaching sessions (HBCH99). In contrast, robot-human interactions in the AURORA project are it unconstrained and unstructured, the children are allowed to interact with the robot in whatever body position they prefer (e.g. lying on the floor, crawling, standing, cf. figure 4, they are also free to chose how they interact with the robot (touching, approaching, watching from a distance, picking it up etc.). Interference is only necessary if the child is about to damage the robot or if the child (by pressing buttons) switches off the robot so that it needs to be restarted. Such conditions are much different from other projects on robot-human interaction which are based on structured and constrained set ups (e.g. KISMET, or the ROBOTA dolls) where the human is expected to interact with the robot while adopting a particular position and orientation towards the robot (e.g. sitting face-to-face in close distance to an interactive robot that is not moving in space). The particular challenges faced in the AURORA project, in the broader context of rehabilitation, together with a more detailed discussion of therapeutical issues involved, is given in (WD99), (DW00).

Theoretical Background and Working Hypotheses

The AURORA project deliberately uses a non-humanoid robot, based on the observation that children with autism prefer a predictable, stable environment and that many people with autism have difficulty interpreting facial expressions and other social cues in social interactions. Consequently, they often avoid social interactions since people appear unpredictable and confusing. Generally, using a robot as a remedial toy takes up the challenge of bridging the gap between the variety and unpredictability of human social behaviour (which often appears frightening to children with autism) and the predictability of repetitive and monotonous behaviour which children with autism prefer and which can be performed by mobile robots (see discussion in (Dau99c)). We hypothesise that a child with autism 1) is sufficiently interested in 'playing' with an interactive autonomous robot as it is used in the AURORA project, 2) the robot can engage the child in interactions which demonstrate important aspects of human-human interaction (e.g. eye-contact, turn-taking, imitation games), and 3) (as a long term therapeutic goal), while slowly increasing the robot's behaviour repertoire and the unpredictability of its actions and reactions, the robot can be used to guide the children towards more realistic and 'complex' forms of social interactions resembling human-human interaction. This approach is based on two areas of theoretical work, namely mindreading and interaction dynamics. These issues and their implications for the AURORA project are described in the following two sections.

Mindreading

Generally, humans are from an early age on attracted to self-propelled objects which are moving autonomously and seemingly with 'intention' (Dau97). In (PP95b) a theory of human social competence is presented that consists of three units: the first unit (intentional system) identifies self- propelled movements in space and interprets them as intentional, engaged in goal-directed behaviour, such as escaping from confinement, making contact with another intentional object, overcoming gravity (e.g. seeking to climb a hill). Animate and inanimate objects are distinguished since only animate objects can move both in space and time without the influence of other objects. Movement in place is interpreted as animate but not intentional. The second unit is the social system which specifies the changes that the intentional objects undergo. It allows to interpret relations e.g. as possession or group membership. The third unit is the theory of mind system, which outputs explanation, states of mind, perception, desire, belief, and its variations. These mental states are used to explain the actions. Premack and Premack's theory of human social competence shows great similarity with Baron-Cohen's suggestion of four mechanisms underlying the human mindreading system (BC95). The first mechanism is the intentionality detector that interprets motion stimuli (stimuli with self- propulsion and direction) in terms of the mental states of goal and desire. These primitive mental states are basic since they allow making sense of universal movements of all animals, namely approach and avoidance, independent of the form or shape of the animal. The ID mechanism works through vision, touch and audition and interprets anything that moves with self-propelled motion or produces a non-random sound as an object with goals and desires. The second mechanism as part of Baron-Cohen's mindreading system is the eye-direction detector (EDD) which works only through vision. The EDD detects the presence of eye-like stimuli, detects the direction of eyes, and interprets gaze as seeing (attribution of perceptual states). This mechanism allows interpreting stimuli in terms of what an agent sees. ID and EDD represent dyadic relations (relations between two objects, agent and object or agent and self) such as 'Agent X wants Y' or 'Agent X sees Y', however they not allow to establish the link between what another agent sees and wants and what the self sees and wants. Sharing perceptions and beliefs is beyond the 'autistic universe', it requires the additional mechanisms SAM (shared-attention-mechanism, allows to build triadic representations: relations between an agent, the self, and a third object) and TOM (theory-of-mind mechanism). ID, EDD, SAM and TOM make up a fully developed human mindreading system as it exists in biologically normal children above the age of four. In normal development, from birth to about 9 months a child can only build dyadic representations based on ID and basic functions of EDD. From about 9 to 18 months SAM comes on board and allows triadic representations that make joint attention possible. SAM links EDD and ID, so that eye direction can be read in terms of basic mental states. From about 18 to 48 months TOM comes on board, triggered by SAM. The arrival of TOM is visible e.g. through pretend play. Note, that earlier mechanisms are not replaced by newer ones, they still continue to function. According to Simon- Baron's analysis children with autism possess ID and EDD. TOM is missing in all children with autism while some of them possess SAM. Referring to this theoretical framework, the working hypotheses (section 2.4) studied in the AURORA project clearly address the ID and EDD mechanisms. In the same way as biologically normal children above 4 years of age detect, are attracted to, and interpret autonomous, self-propelled objects such as robots as 'social agents', we hypothesise that children with autism can accept a mobile robot as a social agent.

Figure 4: An autistic boy playing with the Labo-1 mobile robot which was kindly donated by *Applied AI Systems* Inc. The child is not afraid to let the robot come physically very close to his body, including the face.

Figure 5: The child frequently 'reaches out' to the robot, 'testing' its front sensors and eliciting the robot's response to approach or avoid. After 20 minutes the teacher ended the interaction since the child had to go back to class.

Interaction Dynamics

The second strand of theories which the AURORA project is influenced by concerns interaction dynamics between babies and their caretakers as studied in developmental psychology, e.g. (Bul79), (UBKV89), (Mel96), (MM99). A more detailed account of these issues in the general context of robot-human interaction, and their relevance in the AURORA project is given in (DW00), we can only present a brief summary here. Infants seem to detect specific temporal and structural aspects of infant-caretaker interaction dynamics. It is suggested that turn-taking and imitation games allow the infant 1) to identify people as opposed to other objects, and 2) to use the like-me-test in order to distinguish between different persons. Motivated by this research we suggested a conceptual framework in order to classify different and increasingly complex dynamics in robot-human interactions (DW00). Within this framework, robot-human interactions in the AURORA project are designed where synchronisation of movements, temporal coordination, and the emergence of imitation games are used as important mechanisms for making 'social contact' between the robot and the child. It is hoped that such an approach which focuses on interaction dynamics rather than cognitive reasoning mechanisms can incrementally facilitate and strengthen temporal aspects which are so fundamental to the development of social competence and the ability to socially interact with people (cf. (Hal83).

AURORA: Preliminary Results[9]

Initial trials in the AURORA project stressed the individual nature of the specific needs of children with autism, but they also showed that most children responded very well and with great interest to the autonomous robot, see figures 4, 5. In a recent series of comparative trials where the children were playing with the robot (condition 1) and also (separately) with a passive non-robotic toy (condition 2) children showed greater interest in interactions with the robot than with the ''inanimate' toy (quantitative data will be published in a forthcoming publication by Werry and Dautenhahn). Also, children often showed increased interest in the front part of the robot where the pyro-electric sensor is attached, a sensor with strongly eye-like features (eye- like shape, located at the distal end of the robot's preferred direction of movement, prominent position raised above the chassis, direction of the sensor changing according to 'gaze'). These observations seem to confirm our hypothesis that interactions in the AURORA project can successfully built on mechanisms of intentionality detection (ID mechanism) and eye-direction- detection (EDD mechanism). Please note that mobile robots are only seen as potentially one form of therapy, which might complement other forms of therapies (see review in (DW00)). An interesting line for future research is to study the application of virtual environments for children with autism, as discussed in (Dau00a).

A particular problem we encounter in the AURORA project is that (with few exceptions) we cannot ask our subjects, they do not give verbal feedback, techniques like interviews or questionnaires are impossible. This puts particular emphasis on the analysis of behaviour and interaction. I believe that the field of socially intelligent agents has a huge potential in education, therapy and rehabilitation. However, new design and evaluation techniques and methodologies need to be developed (cf. (MJL00), (MAD$^+$00).

[9] For more recent results and evaluations see (DWR$^+$02) and K. Dautenhahn, I. Werry (2002) A Quantitative Technique for Analysing Robot-Human Interactions. Proc. IROS2002, Lausanne, 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems

Empathy

In (Dau97) (see also (Dau99a)) I discussed empathy as a fundamental, experiential mechanism with humans use to bond and understand each other. Also, empathy can be considered as a means of social learning via bonding with other people. According to Wispe empathy is a way of 'knowing', as opposed to 'relating' which occurs in sympathy (Wis86). Brothers considers empathy (Bro89) a biological phenomenon, an 'emotional communication' system that the human *social brain* seems to be specialised in (Bro97).

Inspired by autism research and Barrett-Lennard's cyclic/phasic model of empathy, (BL81), (BL93), I suggested in (Dau97) and (Dau99a) to distinguish between two different mechanisms: a) *empathic resonance*, an immediate, direct way of re-experiencing, and b) *biographical reconstruction*, namely reading another persons mind by re-constructing the other's autobiographical context (who that person is, where he comes from, what the relationship is with oneself, what behaviour might be expected etc.). Barrett-Lennard's empathy cycle is a process between two people, involving expressing and receiving empathy.

Recent experiments on empathic accuracy (the ability to read and understand reliably another person's intentions, beliefs, etc.), as well as neurophysiological experiments with monkeys point towards an exciting possibility how empathic resonance might actually be grounded in biological mechanisms: Neurons were found in area F5 of the monkey brain that discharge when the monkey grasps or manipulates objects, but also when the monkey observes an experimenter making a similar gesture ((GFFR96), (RFGF98)). Arbib (Arb02) speculates that all primates (including humans) might share the mirror system as a neurobiological mechanism underlying imitation (note that imitation can be shown for humans and other apes, but is difficult to confirm for monkeys, who are nevertheless good social learners, (VF02)). Some researchers

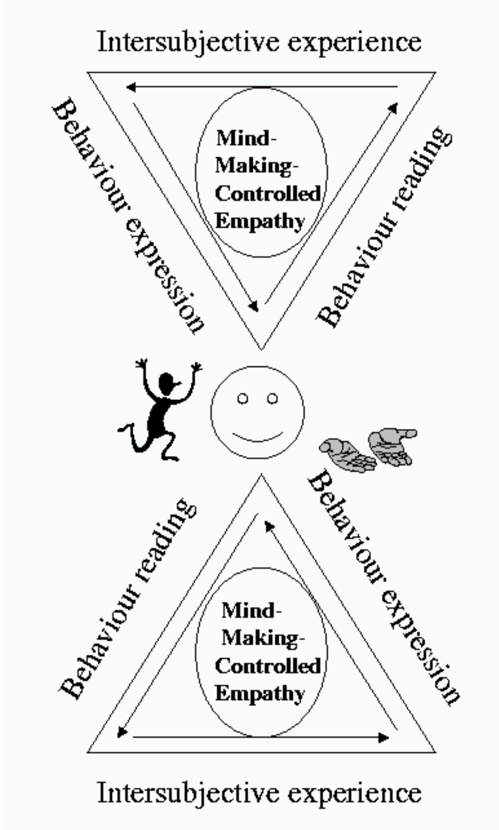

Figure 6: A Theory of empathy based on issues discussed in this paper, see text for explanation.

even suggest that the mirror system in F5 (analogous to Broca's area in humans, important for language), is 'grounding' language in gestures and body language (Arb02), (RA98). Although the findings of neurophysiological studies in monkeys need to be confirmed for humans and further understood for all primates and non-primate animals, it is suggested that mirror neurons could be the basis for a simulation theory of empathy (GG98). Previously, in discussions on how people ascribe mental states to themselves and others, a *simulation theory* was opposed to a *theory theory* (Gol92), (Gor92). Supporters of the simulation theory favour a process of 'putting oneself in the other's place', as opposed to (detached) reasoning about other's beliefs, emotions etc. The importance of the mirror system in this context is that (in support of the simulation theory) it could be nature's solution - at least in some primates - to solving the correspondence problem for empathy and creating *intersubjective experience* by creating a common shared context and shared *understanding* of actions and affordances. Figure 6 shows mechanisms and processes which were discussed above and how they might fit into a theory of empathy. Here, two persons are linked via a common social 'currency', namely facial expressions, body language, gestures, imitation games, interactions dynamics, spatial-temporal dynamics as they are are studied in proxemics (the study of human's perception and use of space, cf. (Hal68) (Hal83)), etc. These are important for *automatic empathy* that creates *intersubjective experience* and *physiological synchrony* (LR97). This is accompanied by a cognitive, *controlled mechanism of empathy*, a more deliberative inference making (HW97), what we called biographical reconstruction. The theory sketched in figure 6 needs to be confirmed by the discovery mirror neurons e.g. for facial expressions and other gestures, as speculated in (Bro97), p. 78. Simon Baron-Cohen's theory of mindreading nicely fits in this framework, as well as Mitchell and Hamm's discussion of behaviour reading, see above. ID and EDD might (in normally developed humans) play a strong role in behaviour reading (although they do not seem to be necessary, since empathy does not rely on the visual channel alone, cf. blind people, or Mitchell and Hamm's study of behaviour reading with narratives (MH97)). TOM might be part of controlled empathy, reasoning about another person's beliefs, desires, goals, emotions etc. and biographical reconstruction.

Empathy does not only occur in face-to-face contact with another person, it can also be evoked by reading a book or watching a movie, i.e. without any feedback from the character/person we might empathise with. If I empathise with a human being (whether real, enacted, fictional or imagined) empathy is nevertheless based on my assumption that the other human is to some extent 'like me'. An important challenge is then to create empathic relationships with non-human artifacts. Here, as shown in figure 7 we cannot assume an 'understanding' from the artifact. However, our own behaviour-reading and expression mechanisms still work, and non-biological socially intelligent agents could exploit this, see (Dau95), (BDH98), (BS00), (BA00), and other research projects on social robots described in this volume and elsewhere.

Quo Vadis?

I argued in this paper that Socially Intelligent Agents (SIA) research, although strongly linked to software and robotic engineering, goes beyond a software engineering paradigm: it can potentially serve as a paradigm for a *science of social minds*. This paper gave some indications of a few research questions that I believe are important. A systematic and experimental investigation of human social minds and the way humans perceive the social world can result in truly

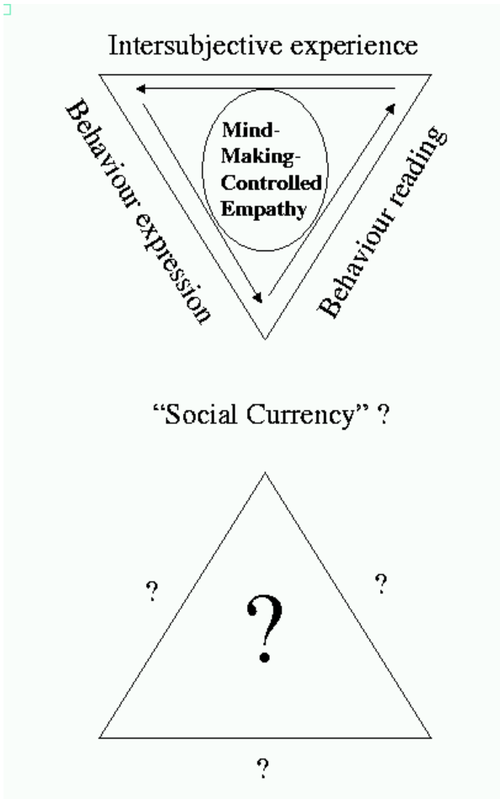

Figure 7: Empathising with agents?

social artifacts, socially intelligent agents that are integrated in human society, e.g. social robots that meet the cognitive and social needs of humans[10]. Such social agents might become more and more similar to us, in ways which could even make it difficult to distinguish between them and human beings, cf. (Fon97), (Fon00), (Dau00c).

> "*Once there was a robot called Jig Jag and Jig Jag lived in the countryside. One day Jig Jags lights started to flash, that meant that the robot had an idea. "I think I will go for a walk", so Jig Jag went into a field with some sheep in it and the silly robot tried to talk to the sheep, "Silly, silly, Jig Jag". Next Jig Jag saw some cows in the next field, so silly Jig Jag tried to talk to the cows! After that Jig Jag went to the shops, he wanted to buy some bolts and oil. So Jig Jag went into the hardware shop, but poor Jig Jag set the alarm off. So Jig Jag went into another hardware store across the road. So the robot tried to get into the shop but again Jig Jag set the alarm off. So poor Jig Jag had to go home empty handed.*" (Lauren - 8 years old, (BD99)).

[10]For a recent discussion on roles of robots in human society see K. Dautenhahn: Roles and Functions of Robots in Human Society - Implications from Research in Autism Therapy, accepted for publication in *Robotica*, special issue on Biological Robotics, guest-editor: David McFarland, Cambridge University Press.

Acknowledgements

I like to thank Katherine Bumby and Iain Werry whose work is summarised in this paper and helped me to develop my ideas. The AURORA project is supported by an EPSRC grant (GR/M62648).

References

Michael Arbib. The mirror system, imitation, and the evolution of language. In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, 2002.

Alan H. Bond and Matt Ashton. The robot as person. Workshop on Interactive Robotics and Entertainment (WIRE-2000), Pittsburgh, April 2000, 2000.

Steve Benford, Benjamin B. Bederson, Karl-Petter Akesson, Victor Bayon, Allison Druin, Par Hansson, Juan Pablo Hourcade, Rob Ingram, Helen Neale, Claire O'Malley, Kristian T. Simsarian, Danae Stanton, Yngve Sundbald, and Gustav Taxen. Designing storytelling technologies to encourage collaboration between young children. *Proc. CHI 2000*, April 1-6, 2000, The Hague, The Netherlands, 2000.

Simon Baron-Cohen. *Mindblindness: An Essay on Autism and Theory of Mind*. A Bradford Book, The MIT Press, Cambridge, London, 1995.

M. Umaschi Bers and J. Cassell. Children as designers of interactive storytellers "let me tell you a story about myself". In Kerstin Dautenhahn, editor, *Human Cognition and Social Agent Technology*, chapter 16, pages 61–83. John Benjamins Publishing Company, 2000.

Robert A. Barton and Robin I.M. Dunbar. Evolution of the social brain. In Andrew Whiten and Richard W. Byrne, editors, *Machiavellian Intelligence II Extensions and Evaluations*, chapter 9, pages 240–263. Cambridge University Press, 1997.

K. Bumby and K. Dautenhahn. Investigating children's attitudes towards robots: A case study. In *Proc. CT99, The Third International Cognitive Technology Conference, August, San Francisco, available at http://www.cogtech.org/CT99*, pages 391–410, 1999.

A. Billard, K. Dautenhahn, and G. Hayes. Experiments on human-robot communication with robota, an imitative learning and communication doll robot. Technical Report CPM-98-38, Centre for Policy Modelling, Manchester Metropolitan University, UK, 1998.

Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Sys-*

tems. Oxford University Press, New York, Oxford, 1999.

Aaron F. Bobick, Stephen S. Intille, James W. Davis, Freedom Baird, Claudio S. Pinhanez, Lee W. Campbell, Yuri A. Ivanov, Arjan Schtte, , and Andrew Wilson. The kidsroom: A perceptually-based interactive and immersive story environment. *Presence*, 8(4):369–393, August 1999.

Godfrey T. Barrett-Lennard. The empathy cycle: refinement of a nuclear concept. *Journal of Counseling Psychology*, 28(2):91–100, 1981.

Godfrey T. Barrett-Lennard. The phases and focus of empathy. *British Journal of Medical Psychology*, 66:3–14, 1993.

Susan Blackmore. *The Meme Machine*. Oxford University Press, 1999.

Katherine H. Blocher. Affective Social Quest (ASQ). Teaching emotion recognition with interactive media and wireless expressive toys. Master's Thesis for Master of Science in Media Technology Massachusetts Institute of Technology, MIT, USA, 1999.

D. Berdichevsky and E. Neunschwander. Toward and ethics of persuasive technology. *Communications of the ACM*, 42(5):51–58, 1999.

Valentin Braitenberg. *Vehicles: Experiments in Synthetic Psychology*. MIT Press, Cambridge, 1984.

Leslie Brothers. A biological perspective on empathy. *American Journal of Psychiatry*, 146(1):10–19, 1989.

Leslie Brothers. *Friday's footprint: how society shapes the human mind*. Oxford University Press, New York, Oxford, 1997.

Jerome Bruner. The Narrative Construction of Reality. *Critical Inquiry*, 18(1):1–21, 1991.

Cynthia Breazeal and Brian Scassellati. Infant-like social interactions between a robot and a human caretake. *Adaptive Behavior*, 8(1):49–74, 2000.

M. Bullowa. *Before Speech*. Cambridge University Press, 1979.

Richard W. Byrne and Andrew Whiten. Machiavellian intelligence. In Andrew Whiten and Richard W. Byrne, editors, *Machiavellian Intelligence II Extensions and Evaluations*, chapter 1, pages 1–23. Cambridge University Press, 1997.

R. W. Byrne. Machiavellian intelligence. *Evolutionary Anthropology*, 5:172–180, 1997.

Steven Coles and Kerstin Dautenhahn. A robotic story-teller. Proc. *SIRS2000*, 8th Symposium on Intelligent Robotic Systems, The University of Reading, England, 18-20 July 2000, pp. 393-401, 2000.

Kerstin Dautenhahn. Getting to know each other – artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

Kerstin Dautenhahn. I could be you – the phenomenological dimension of social understanding. *Cybernetics and Systems*, 25(8):417–453, 1997.

Kerstin Dautenhahn. The art of designing socially intelligent agents: science, fiction and the human in the loop. *Applied Artificial Intelligence Journal, Special Issue on Socially Intelligent Agents*, 12(7-8):573–617, 1998.

Kerstin Dautenhahn. Embodiment and interaction in socially intelligent life-like agents. In C. L. Nehaniv, editor, *Computation for Metaphors, Analogy and Agents*, pages 102–142. Springer Lecture Notes in Artificial Intelligence, Volume 1562, 1999.

Kerstin Dautenhahn. The lemur's tale - story-telling in primates and other socially intelligent agents. *Proc. Narrative Intelligence, AAAI Fall Symposium 1999*, AAAI Press, Technical Report FS-99-01, pp. 59-66, 1999.

Kerstin Dautenhahn. Robots as social actors: Aurora and the case of autism. In *Proc. CT99, The Third International Cognitive Technology Conference, August, San Francisco*, pages 359–374, 1999.

Kerstin Dautenhahn. Design issues on interactive environments for children with autism. In Paul Sharkey, Antonio Cesarani, Luigi Pugnetti, and Albert Rizzo, editors, *Proc. The 3rd International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2000, 23-25 September 2000, Alghero, Sardinia, Italy*, pages 153–161, 2000.

Kerstin Dautenhahn. Evolvability, culture and the primate social brain. In C. L. Nehaniv, editor, *Proceedings of the Evolvability Workshop at the Seventh International Conference on the Simulation and Synthesis of Living Systems (Artificial Life VII)*, pages 23–26. 2000.

Kerstin Dautenhahn. Reverse engineering of societies - a biological perspective. Proc. AISB Symposium *Starting from Society - the application of social analogies to computational systems*, a symposium at AISB-00, 17th-20th April 2000, University of Birmingham, England, pp. 21-31, 2000.

Richard Dawkins. *The Selfish Gene*. Oxford University Press, 1976.

Kerstin Dautenhahn and Steven Coles. Robotic storytellers: A bottom-up computational framework for the study of narrative intelligence in autonomous agents. *Proc. Narrative and Interactive Learning Environments* Edinburgh, Scotland 30th August - 1st September 2000, pp. 51-59, 2000.

D. C. Dennett. Intentional systems. *Journal of Philosophy*, 68:87–106, 1971.

Daniel C. Dennett. *The intentional stance*. MIT Press, 1987.

Daniel C. Dennett. The origins of selves. *Cogito*, 3, pp. 163-73, Autumn 1989. Reprinted in Daniel Kolak and R. Martin, eds., *Self and Identity: Contemporary Philosophical Issues*, Macmillan, 1991, 1989.

Merlin Donald. Precis of origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences*, 16:737–791, 1993.

R. I. M. Dunbar. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16:681–735, 1993.

Frans B.M. de Waal. Cultural primatology comes of age. *Nature*, 399:635–636, 1999.

Kerstin Dautenhahn and Iain Werry. Issues of robot-human interaction dynamics in the rehabilitation of children with autism. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, and S. W. Wilson, editors, *Proc. From animals to animats 6, The Sixth International Conference on the Simulation of Adaptive Behavior (SAB2000)*, pages 519–528, 2000.

Kerstin Dautenhahn, Iain Werry, John Rae, Paul Dickerson, Penny Stribling, and Bernard Ogden. Robotic playmates – analysing interactive competencies of children with autism playing with a mobile robot. In K. Dautenhahn, A. Bond, L. Ca namero, and B. Edmonds, editors, *Socially Intelligent Agents – Creating Relationships with Computers and Robots*, pages 117–124. Kluwer Academic Publishers, 2002.

B. Edmonds and K. Dautenhahn. The contribution of society to the construction of individual intelligence. Techical Report CPM-98-42, Centre for Policy Modelling, Manchester Metropolitan University, UK, 1998.

T. J. Eddy, G. G. Gallup, and D. J. Povinelli. Attribution of cognitive states to animals: Anthropomorphism in comparative perspective. *Journal of Social Issue*, 49(1):87–101, 1993.

Stan Franklin and Art Graesser. Is it an agent, or just a program?: A taxonomy for autonomous agent. In *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, published as Intelligent Agents III*, pages 21–35. Springer-Verlag, 1997.

B. J. Fogg. Introduction: Persuasive technologies. *Communications of the ACM*, 42(5):27–29, 1999.

Leonard N. Foner. Entertaining agents: a sociological case study. In W. Lewis Johnson, editor, *Proc. of the First International Conference on Autonomous Agents, Marina del Rey, CA, USA, February 5-8*, pages 122–129, 1997.

Leonard N. Foner. Are we having fun yet? using social agents in social domains. In Kerstin Dautenhahn, editor, *Human Cognition and Social Agent Technology*, chapter 12, pages 323–348. John Benjamins Publishing Company, 2000.

Stan Franklin. Autonomous agents as embodied ai. *Cybernetics and Systems*, 28(6), 1997.

Jennifer W. Glos and Justine Cassell. Rosebud: A place for interaction between memory, story, and self. In Jonathon P. Marsh, Chrystopher L. Nehaniv, and Barbara Gorayska, editors, *Proceedings of the Second International Conference on Cognitive Technology*, pages 88–97. IEEE Computer Society Press, 1997.

V. Gallese, L. Fadiga, L. Fogassi, and G. Rizzolatti. Action recognition in the premotor cortex. *Brain*, 119:593–609, 1996.

V. Gallese and A. Goldman. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12):493–501, 1998.

Gerd Gigerenzer. The modularity of social intelligence. In Andrew Whiten and Richard W. Byrne, editors, *Machiavellian Intelligence II Extensions and Evaluations*, chapter 10, pages 264–288. Cambridge University Press, 1997.

Alvin I. Goldman. In defense of the simulation theory. *Mind and Language*, 7(1-2):104–119, 1992.

Robert M. Gordon. The simulation theory: objections and misconceptions. *Mind and Language*, 7(1-2):11–33, 1992.

Temple Grandin. *Thinking in pictures*. Doubleday Publisher, New York, London, Toronto, Sydney, Auckland, 1995.

T. Grandin and M. M. Scariano. *Emergence: Labeled autistic*. Warner Books, 1996.

E. T. Hall. Proxemics. *Current Anthropology*, 9(2-3):83–95, 1968.

E. T. Hall. *The Dance of Life: The Other Dimension of Time*. Anchor Books, Doubleday, 1983.

Patricia Howlin, Simon Baron-Cohen, and Julie Hadwin. *Teaching Children with Autism to Mind-Read*. John Wiley and Sons, 1999.

F. Heider and M. Simmel. An experimental study of apparent behavior. *American Journal of Psychology*, 57:243–259, 1944.

Michael A. Huffman. Acquisition of innovative cultural behaviors in nonhuman primates: a case study of stone handling, a socially transmitted behaviour in japanese macaques. In Cecilia M. Heyes and Bennett G. Galef Jr., editors, *Social learning in animals*, chapter 13, pages 267–289. Academic Press, 1996.

S. D. Hodges and D. M. Wegner. Automatic and controlled empathy. In W. Ickes, editor, *Empathic accuracy*, pages 311–339. The Guildford Press, 1997.

N. R. Jennings, K. Sycara, and M. Wooldrige. A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1:7–38, 1998.

Marc Kirschner and John Gerhart. Evolvability. *Proc. Natl. Acad. Sci*, 95:8420–8427, 1998.

R. W. Levenson and A. M. Ruef. Physiological aspects of emotional knowledge and rapport. In W. Ickes, editor, *Empathic accuracy*, pages 44–72. The Guildford Press, 1997.

Jaime Montemayor, Houman Alborzi, Allison Druin, Jim Hendler, Deborah Pollack, Jessica Porteous, Lisa Sherman, Asmara Afework, Jack Best, Joe Hammer, Aleander Kriskal, Abigail Lal, Thomas Plaisant Schwenn, Lauren Sumida, and Rebecca Wagner. From pets to storykit: Creating new technology with an intergenerational design team. *Proceedings 2000 Workshop on Interactive Robotics and Entertainment (WIRE-2000) in cooperation with AAAI, April 30 - May 1, 2000, The Robotics Institute Carnegie Mellon University Pittsburgh, Pennsylvania, USA*, 2000.

Michael Mateas. Computational subjectivity in virtual world avatars. In *Socially Intelligent Agents*, pages 43–45. AAAI Press, Technical report FS-97-02, 1997.

F. Michaud, A. Clavet, G. Lachiver, and M. Lucas. Designing toy robots to help autistic children - an open design project for electrical and computer engineering education. *Proc. American Society for Engineering Education*, 2000.

A. Meltzoff. The human infant as imitative generalist: a 20-year progress report on infant imitation with implications for comparative psychology. In B. G. Galef and C. M. Heyes, editors, *Social Learning in Animals: the Roots of Culture*, pages 347–370. Academic Press, New York, 1996.

R. W. Mitchell and M. Hamm. The interpretation of animal psychology: Anthropomorphism or behavior reading? *Behaviour*, 134:173–204, 1997.

Steven Mithen. *The prehistory of the mind*. Thames and Hudson Ltd, London, England, 1996.

Stacy C. Marsella, W. Lewis Johnson, and C. LaBore. Interactive pedagogical drama. In *Proceedings of the Fourth International Conference on Autonomous Agents, June 3-7, Barcelona, Spain*, pages 301–308. ACM Press, 2000.

F. Michaud, P. Lepage, J.-D. Leroux, M. Clarke, F. Belanger, Y Brosseau, and D. Neu. Mobile robotic toys for autistic children. Proc. International Symposium on Robotics, Montral, May 2000, 2000.

A. N. Meltzoff and M. K. Moore. Persons and representation: why infant imitation is important for theories of human development. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 9–35. Cambridge University Press, 1999.

I. Machado, C. Martinho, and A. Paiva. Once upon a time. *Proc. Narrative Intelligence, AAAI Fall Symposium 1999*, AAAI Press, Technical Report FS-99-01, pp. 115-119, 1999.

NAS. http:www.oneworld.org/autism_uk/index.html. Last referenced on 14th of February, 2000, 2000.

Chrystopher L. Nehaniv and Kerstin Dautenhahn. Living with socially intelligent agents: a cognitive technology view. In Kerstin Dautenhahn, editor, *Human Cognition and Social Agent Technology*, chapter 16, pages 415–426. John Benjamins Publishing Company, 2000.

Seymour Papert. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, New York, 1980.

Daniel J. Povinelli and Todd M. Preuss. Theory of mind: evolutionary history of a cognitive specialization. *Trends in Cognitive Neurosciences*, 18:418–424, 1995.

David Premack and Ann James Premack. Origins of human social competence. In Michael S. Gazzaniga, editor, *The cognitive neurosciences*, pages 205–218. A Bradford Book, The MIT Press, 1995.

D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4:515–526, 1978.

T. Quick, K. Dautenhahn, C. Nehaniv, and G. Roberts. The essence of embodiment: A framework for understanding and exploiting structural coupling between system and environment. Proc. CASYS'99, Third International Conference on Computing Anticipatory Systems, HEC, Lige, Belgium, August 9 -14, 1999.

T. Quick, K. Dautenhahn, C. Nehaniv, and G. Roberts. On bots and bacteria: Ontology independent embodiment. Proc. ECAL99, Fifth European Conference on Artificial Life, Switzerland, September, pp. 339-343, 1999.

T. Quick, K. Dautenhahn, C. Nehaniv, and G. Roberts. Understanding embodiment, system-environment coupling and the emergence of adaptive behaviour. In: A. Drogoul, J.-A. Meyer (eds.) Intelligence artificielle située, Hermes Science Publications, Paris, 1999, pp. 13-31, 1999.

G. Rizzolatti and M. A. Arbib. Language within our grasp. *Trends in Neurosciences*, 21(5):188–194, 1998.

Mitchel Resnick. LEGO, LOGO, and Life. In C. G. Langton, editor, *Proc. of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Los Alamos, New Mexico, September 1987*, pages 397–406, 1989.

G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognition and Brain Research*, 3:131–141, 1998.

Simon M. Reader and Kevin N. Laland. Do animals have memes? *Journal of Memetics - Evolutionary Models of Information Transmission*, 3(2):., 1999.

Stephen John Read and Lynn Carol Miller. Stories are fundamental to meaning and memory: for social creatures, could it be otherwise? In Robert S. Wyer, editor, *Knowledge and Memory: the Real Story*, chapter 7, pages 139–152. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1995.

B. Reeves and C. Nass. *The Media Equation*. Cambridge University Press, 1996.

Robert Jay Russell. *The Lemurs' Legacy. The Evolution of Power, Sex, and Love.* G.P. Putnam's Sons, New York, 1993.

Susanne Schäfer, editor. *Sterne, Äpfel und rundes Glas. Mein Leben mit Autismus.* Verlag Freies Geistesleben und Urachhaus GmbH, Stuttgart, Germany, 1997.

Phoebe Sengers. Do the thing right: an architecture for action-expression. In Katia P. Sycara and Michael Wooldrige, editors, *Proc. of the Second International Conference on Autonomous Agents, Minneapolis/St. Paul, USA, May 9-13*, pages 24–31, 1998.

Phoebe Sengers. Narrative intelligence. In: *Human Cognition and Social Agent Technology*, Ed. Kerstin Dautenhahn, John Benjamins Publishing Company, 2000.

Paul W. Sherman, Jennifer U.M. Jarvis, and Richard D. Alexander, editors. *The Biology of the Naked Mole-Rat.* Princeton University Press, Princeton, N.J, 1991.

S. Tseng and B. J. Fogg. Credibility and computing technology. *Communications of the ACM*, 42(5):39–44, 1999.

M. Tomasello, A.C. Kruger, and H.H. Ratner. Cultural learning. *Behavioral and Brain Sciences*, 16(3):495–552, 1993.

I. Uzgiris, J. Benson, J. Kruper, and M. Vasek. Establishing action-environment correspondences: Contextual influences on imitative interactions between mothers and infants. In J. Lockman and N. Hazen, editors, *Action in Social Context*, pages 103–127. New York: Plenum, 1989.

Elisabetta Visalberghi and Dorothy Fragaszy. Do monkeys ape? - ten years after. In K. Dautenhahn and C. L. Nehaniv, editors, *Imitation in Animals and Artifacts*. MIT Press, 2002.

Stuart N. K. Watt. Seeing things as people: anthropomorphism and common-sense psychology. Unpublished PhD thesis, Department of Psychology, The Open University, 1997.

A. Whiten and R. W. Byrne. The machiavellian intelligence hypotheses: editorial. In R. W. Byrne and A. Whiten, editors, *Machiavellian intelligence*, chapter 1. Clarendon Press, 1988.

Iain Werry and Kerstin Dautenhahn. Applying robot technology to the rehabilitation of autistic children. Proc. SIRS99, 7th International Symposium on Intelligent Robotic Systems '99, 1999.

S. Weir and R. Emanuel. Using LOGO to catalyse communication in an autistic child. Technical report, DAI Research Report No. 15, University of Edinburgh, 1976.

A. Whiten, J. Goodall, W.C. McGrew, T. Nishida, V. Reynolds, Y. Sugiyama, C.E.G. Tutin, R.W. Wrangham, and C. Boesch. Cultures in chimpanzees. *Nature*, 399:682–685, 1999.

Andrew Whiten. *Natural Theories of Mind*. Basil Blackwell, 1991.

Lauren Wispe. The distinction between sympathy and empathy: to call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, 50(2):314–321, 1986.

M. Wooldridge, N. R. Jennings, and D. Kinny. The gaia methodology for agent-oriented analysis and design. *Journal of Autonomous Agents and Multi-Agent Systems*, 3(3):285–312, 2000.

About the Tutors

Kerstin Dautenhahn is Reader in Artificial Intelligence in the Department of Computer Science at University of Hertfordshire. She is editor of “Human Cognition and Social Agents Technology” (John Benjamins Publishing Company, 2000) and “Imitation in Animals and Artefacts” (MIT Press, to appear in April 2002.) Previously, she gave a half-day tutorial "On Minds and Agents: Social Intelligence in Animals and Artifacts" at Autonomous Agents 2000, Fourth International Conference on AUTONOMOUS AGENTS (Agents 2000), Barcelona, Spain, June 2000, and a two-hour tutorial “Socially Intelligent Agents - From Animals to Animats” within the Joint Tutorial Programme of SAB2000: 6^{th} Int'l Conf. on the Simulation of Adaptive Behavior and PPSN2000: 6^{th} Int'l Conf. on Parallel Problem Solving From Nature, Paris, France, Sunday 17^{th} September 2000.

Home page: `homepages.feis.herts.ac.uk/~comqkd/`

Bruce Edmonds is Research Fellow at the Centre for Policy Modelling, a research unit that specialises on social simulation. His areas of research include: the methodology of social simulation, evolutionary computation, and context-dependency. He is a co-editor of the Journal of Memetics - Evolutionary Models of Information Transmission. He was part of the team that developed SDML a development environment for agent-based social simulation. He has recently co-edited special issues on the topics of "Computational Memetics" and "Context in context".

Home page: `bruce.edmonds.name`

Some Relevant Web Pages

This is in no particular order – only sites with substantive information listed, i.e. not merely tables of contents.

The Socially Intelligent Agents Webpage:

`http://homepages.feis.herts.ac.uk/~comqkd/aaai-social.html`

The Journal of Artificial Societies and Social Simulation (JASSS)

`http://www.soc.surrey.ac.uk/JASSS/JASSS.html`

The Journal of Memetics – Evolutionary Models of Information Transmission

`http://jom-emit.cfpm.org/`

Social Adeptness in Agents

`http://www.iit.nrc.ca/~steve/SocAdept/`

The Emotion Forum

`http://www.iiia.csic.es/~lola/emotion.html`

The COG Project

`http://www.ai.mit.edu/projects/cog/cog.html`

Emotional Agents

`http://www.cs.tamu.edu/people/magys/projects/em.html`

Simulation/Multi-Agent Evaluation

`http://www.isi.edu/~galk/Eval/`

Online Proceedings of the 1999 Workshop Intelligent Virtual Agents

`http://www.salford.ac.uk/cve/va99/on-line99.htm`

The Social Web Project

`http://orgwis.gmd.de/projects/SocialWeb/`

Papers from Workshop on Socially Situated Intelligence

`http://bruce.edmonds.name/ssi/`

Computational Analysis of Social and Organizational Sustems

`http://www.casos.ece.cmu.edu/home_frame.html`

Papers from the workshop on “Starting from Society”

`http://bruce.edmonds.name/sfs/`

Papers from the special issue on the same topic

`http://jasss.soc.surrey.ac.uk/4/1/contents.html`

Intelligent Agents Repository

`http://www-cia.mty.itesm.mx/~lgarrido/Repositories/IA/agents.html`

Multi-Agent Research Group at USC

`http://www.isi.edu/teamcore/`

Decision Making in Social Systems

`http://www.ecs.soton.ac.uk/~nrj/soc-rat.html`

A Bibliography of (somewhat) Relevant Papers

Akiyama, E. and K. Kaneko (1996). Evolution of Cooperation, Differentiation, Complexity, and Diversity in an Iterated Three-person Game. *Artificial Life* **2**: 293-304.

Akiyama, E. and K. Kaneko (1999). Dynamical Systems Game (the Lumberjacks' Dilemma model). In. *Advances in Artificial Life, Proceedings*. **1674:** 560-565.

Aknine, S. (1999). Exploiting social reasoning of open multi-agent systems to enhance cooperation in hospitals. In. *Artificial Intelligence in Medicine*. **1620:** 133-137.

Andry, P., P. Gaussier, et al. (2001). Learning and communication via imitation: An autonomous robot perspective. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 431-442.

Arafa, Y. and A. Mamdani (2000). *Communicating Affect in a Multi-Agent System*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 1-3.

Arbib, M. (2002). The Mirror System, Imitation, and the Evolution of Language. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Axelrod, R., J. M. Epstein, et al. (1996). Aligning Simulation Models: A Case Study and Results. *Computational and Mathematical Organization Theory* **1**: ?????

Axtel, R. L. and J. M. Epstein (1994). Agent-based Modelling: Understanding Our Creations. *The Bulletin of the Sante Fe Institute* **Winter 1994**: 28-32.

Ayala, G. and A. Hernandez (2000). CASSIEL: Modeling intelligent agents for a lifelong learning environment. In. *Micai 2000: Advances in Artificial Intelligence, Proceedings*. **1793:** 741-747.

Ayala, G. and Y. Yano (1998). A collaborative learning environment based on intelligent agents. *Expert Systems with Applications* **14**(1-2): 129-137.

Aylett, R. (1997). Robots as Socially Intelligent Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 1-3.

Aylett, R. and D. Barnes (1999). Connecting Reflection and Action - a Heterogeneous Multi-Agent Model. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Balch, T. (2000). Hierarchic social entropy: An information theoretic measure of robot group diversity. *Autonomous Robots* **8**(3): 209-237.

Barbeceanu, M. (1998). Coordinating Agents by Role Based Social Constraints and Conversation Plans. In. T. S. a. B. Buchanan. *Proceedings of the Fourteenth National Conference on Artificial Intelligence and the Ninth Innovative Applications of Artificial Intelligence Conference*. Menlo Park, California, AAAI Press American Association for Artificial Intelligence**:** 16--21.

Baylor, A. L. (2000). Beyond butlers: Intelligent agents as mentors. *Journal of Educational Computing Research* **22**(4): 373-382.

Bazzan, A. L. C., R. H. Bordini, et al. (1997). Agents with Moral Sentiments in an Iterated Prisoner's Dilemma Exercise. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 4-6.

Bekkering, H. and W. Prinz (2002). Goal Representations in Imitative Actions. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Bers, M. U. and J. Cassell (1999). Children as Designers of Interactive Storytellers. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Best, M. L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior* **7**(3-4): 289-306.

Billard, A. (2000). *Play, Dreams and Imitation in Robota* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 9-12.

Billard, A. (2002). Imitation: A Means to Enhance Learning of a Synthetic Proto-Language in Autonomous Robots. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Billard, A. (2002). Play, Dreams and Imitation in Robota. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Billard, A. and K. Dautenhahn (1999). Experiments in learning by imitation - Grounding and use of communication in robotic agents. *Adaptive Behavior* **7**(3-4): 415-438.

Billard, A., K. Dautenhahn, et al. (1998). Experiments on human-robot communication with Robota, an imitative learning and communication doll robot. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep38.html

Billeter, J.-B. (1997). Look at What Happens to Telltales and Buffaloes! In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 7-9.

Blocher, K. and R. W. Picard. (2002). Affective Social Quest: Emotion Recognition Therapy of Autistic Children. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Bogner, M., U. Ramamurthy, et al. (1999). "Conscious" and Conceptual Learning In A Socially Situated Agent. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Boivin, M. J. (2001). Feeling humans and social animals: Theological considerations for an evolutionary account of human emotion. *Journal of Psychology and Theology* **29**(4): 314-329.

Boman, M. and H. Verhagen (1998). Social Intelligence as Norm Adaptation. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep39.html

Bond, A. H. (2000). *Modeling Social Relationship by Mutual Commitment, Mutual Control and Mutual Regulation* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 13-17.

Bond, A. H. (2002). Modeling Social Relationship: An Agent Architecture for Voluntary Mutual Control. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Bouhadada, T. and M. T. Laskri (2001). DB-TUTOR: An intelligent tutoring system using a troublemaker companion. In. *Acs/Ieee International Conference on Computer Systems and Applications, Proceedings*: 36-42.

Bouron, T. and A. Collinot (1992). {SAM}: {A} Model to Design Complex Computational Social Agents. In. B. Neumann. *Proceedings of the 10th European Conference on Artificial Intelligence*. Vienna, Austria, John Wiley & Sons: 239--243.

Brazier, F. M. T., P. A. T. van Eck, et al. (2001). Modelling a society of simple agents: From conceptual specification to experimentation. *Applied Intelligence* **14**(2): 161-178.

Breazeal, C. (2002). Designing Sociable Machines: Lessons Learned. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Breazeal, C., A. Edsinger, et al. (2001). Active vision for sociable robots. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 443-453.

Breazeal, C. and B. Scassellati (2002). Issues in Building Robots that Imitate People. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Breazeal, C. and J. Velasquez (1998). Toward teaching a robot `infant' using emotive communication acts. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep40.html

Brent, E. and G. A. Thompson (1999). Sociology: Modeling social interaction with autonomous agents. *Social Science Computer Review* **17**(3): 313-322.

Bridgeman, B. (1992). On the Evolution of Consciousness and Language. *Psycoloqy* **3**: http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?3.15.

Bridgeman, B. (1992). The Social Bootstrapping of Human Consciousness – Reply to Barlow on Bridgeman on Consciousness. **3**: *Psycoloqy* http://www.cogsci.soton.ac.uk/psyc-bin/newpsy?article=3.20.

Brooks, R. A. (1991). Intelligence without Representation. *Artificial Intelligence* **47**(1-3): 139-159.

Bull, L. (1999). On evolving social systems: communication, speciation and symbiogenesis intelligence. *Computational and mathematical organization theory* **5**(3): 281-301.

Burns, T. R. and E. Engdahl (1998). The Social Construction of Consciousness Part 2: Individual Selves, Self-Awareness, and Reflectivity. *Journal of Consciousness Studies* **2**: 166-184.

Byrne Richard, W. and A. Whiten (1988). *Machiavellian intelligence : social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford ; New York, Clarendon Press : Oxford University Press.

Byrne Richard, W. and A. Whiten (1997). *Machiavellian intelligence II : extensions and evaluations*. Cambridge ; New York, NY, USA, Cambridge University Press.

Call, J. and M. Carpenter (2002). Three Sources of Information in Social Learning. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Cañamero, D. and W. V. d. Velde (1997). Socially Emotional: Using Emotions to Ground Social Interaction. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 10-15.

Cañamero, D. and W. V. d. Velde (1999). Emotionally Grounded Social Interaction. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Cañamero, L. (2002). Playing the Emotion Game with Feelix: What Can a LEGO Robot Tell Us about Emotion? In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Canamero, L. and J. Fredslund (2001). I show you how I like you - Can you read it in my face? *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 454-459.

Cao, W., C. G. Bian, et al. (1997). ViSe2 - An application of intelligent agents approaching efficient cooperation. In. *Sixth Scandinavian Conference on Artificial Intelligence*. **40:** 176-186.

Carley, K. (1994). Sociology - Computational Organization Theory. *Social Science Computer Review* **12**(4): 611-624.

Carley, K. and A. Newell (1994). The Nature of the Social Agent. *Journal of Mathematical Sociology* **19**(4): 221-262.

Carley, K. M. (1996). Artificial intelligence within sociology. *Sociological Methods & Research* **25**(1): 3-30.

Carley, K. M. (2002). Computational organization science: A new frontier. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 7257-7262.

Carley, K. M. and D. Svoboda (1996). Modeling Organizational Adaptation as a Simulated Annealing Process,. *Sociological Methods and Research* **25**: 138-168.

Cassell, T. B. a. J. (2000). *How about this Weather? Social Dialogue with Embodied Conversational Agents* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 4-8.

Castelfranchi, C. (1997). Modeling Social Action for {AI} Agents. In. *Proceedings of the 15th International Joint Conference on Artificial Intelligence ({IJCAI}-97)*. San Francisco, Morgan Kaufmann Publishers**:** 1567--1576.

Castelfranchi, C. (1998). Modelling social action for AI agents. *Artificial Intelligence* **103**(1-2): 157-182.

Castelfranchi, C. (2001). Information Agents: The Social Nature of Information and the Role of Trust. In. M. K. a. F. Zambonelli. *Cooperative Information Agents V, Proceedings of the 5th International Workshop CIA 2001*, Springer-Verlag: Heidelberg, Germany**:** 208--210.

Castelfranchi, C., F. de Rosis, et al. (1998). Personality traits and social attitudes in multiagent cooperation. *Applied Artificial Intelligence* **12**(7-8): 649-675.

Castelfranchi, C., F. d. Rosis, et al. (1997). Social Attitudes and Personalities in Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 16-21.

Castelfrianchi, C. (1998). Modeling Social Action for {AI} Agents. *Artificial Intelligence* **103**(1--2): 157--182.

ChaibDraa, B. (1996). Interaction between agents in routine, familiar and unfamiliar situation. *International Journal of Cooperative Information Systems* **5**(1): 1-25.

Chaibdraa, B., R. Mandiau, et al. (1992). Trends in Distributed Artificial-Intelligence. *Artificial Intelligence Review* **6**(1): 35-66.

Chan, T. W. (1995). Social learning systems: An overview. In. *Innovative Adult Learning with Innovative Technologies*. **61:** 101-122.

Chan, T. W., C. W. Hue, et al. (1999). Space of active social learning models. In. *Advanced Research in Computers and Communications in Education, Vol 1 - New Human Abilities for the Networked Society*. **55:** 53-60.

Channon, A. and R. Damper (1998). The Evolutionary Emergence of Socially Intelligent Agents. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep41.html

Chattoe, E. (1998). Just How (Un)realistic are Evolutionary Algorithms as Representations of Social Processes? *Journal of Artificial Societies and Social Simulatio* **1**(3).

Chattoe, E. (2002). Building Empirically Plausible Multi-Agent Systems - A Case Study of Innovation Diffusion. In. K. Dautenhahn, A. Bond, D. Canamero and B. Edmonds. *Socially Intelligent Agents - creating relationships with computers and robots*, Kluwer**:** Chap. 13.

Chou, C. Y., C. J. Lin, et al. (1999). User modeling in simulating learning companions. In. *Artificial Intelligence in Education - Open Learning Environments: New Computational Technologies to Support Learning, Exploration and Collaboration*. **50:** 277-284.

Cliff, D. and J. Bruten (1999). Animat market - Trading interactions as collective social adaptive behavior. *Adaptive Behavior* **7**(3-4): 385-414.

Conati, C. and M. Klawe. (2002). Socially Intelligent Agents in Educational Games. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Conte, R. (1999). Artificial social intelligence: a necessity for agent systems' developments. *Knowledge Engineering Review* **14**(2): 109-118.

Conte, R. (1999). Social intelligence among autonomous agents. *Computational and mathematical organization theory* **5**(3): 203-228.

Conte, R. (2000). Intelligent Social Learning. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep59.html

Conte, R. (2000). The necessity of intelligent agents in social simulation. In. *Applications of Simulation to Social Sciences***:** 19-38.

Conte, R. (2002). Agent-based modeling for understanding social intelligence. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 7189-7190.

Conte, R. and F. Dignum (2001). From social monitoring to normative influence. *JASSS-the Journal of Artificial Societies and Social Simulation* **4**(2): U130-U149.

Conte, R., B. Edmonds, et al. (2001). Sociology and Social Theory in Agent Based Social Simulation: A Symposium. *Computational and Mathematical Organization Theory*.

Conte, R. and M. Paolucci (2001). Intelligent social learning. *JASSS-the Journal of Artificial Societies and Social Simulation* **4**(1): U61-U82.

Cook, i., T. Bickmore, et al. (1999). Autonomous Synthetic Computer Characters as Personal Representatives. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Cooper, B. and P. Brna. (2002). Designing for Interaction: Creating and Evaluating an Empathic Ambience in Computer Integrated Learning Environments. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

D'Amore, M. M. a. R. (2000). *Discovery of Expert Communities and Social Interests* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 119-121.

Dautenhahn, K. (1997). Ants Don't Have Friends--Thoughts on Socially Intelligent Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 22-27.

Dautenhahn, K. (1997). I Could Be You: The phenomenological dimension of social understanding. *Cybernetics and Systems* **28**(417-453).

Dautenhahn, K. (1998). The art of designing socially intelligent agents: Science, fiction, and the human in the loop. *Applied Artificial Intelligence* **12**(7-8): 573-617.

Dautenhahn, K. (1999). Embodiment and Interaction in Socially Intelligent Life-Like Agents. *Lecture Notes in Computer Science* **1562**: 102--142.

Dautenhahn, K. (2000). Making friends with Robots: An artificial life perspective on socially intelligent agents. *International Journal of Psychology* **35**(3-4): 204-204.

Dautenhahn, K. (2000). *Socially Intelligent Agents and the Primate Social Brain --Towards a Science of Social Minds* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 35-51.

Dautenhahn, K. (2001). Socially intelligent agents - The human in the loop. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 345-348.

Dautenhahn, K. and e. al. (2002). Robotic Playmates: Analysing Interactive Competencies of Children with Autism Playing with a Mobile Robot. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Dautenhahn, K. and e. al. (2002). Socially Intelligent Agents: Creating Relationships with Computers and Robots. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Dautenhahn, K. and A. Billard (1999). Bringing up Robots or The Psychology of Socially Intelligent Robots: From Theory to Implementation. In. O. E. a. J. o. r. P. M. u. l. a. J. M. Bradshaw. *Proceedings of the Third Annual Conference on Autonomous Agents ({AGENTS}-99)*. New York, ACM Press**:** 366--367.

Dautenhahn, K., J. Masthoff, et al. (1997). *Socially intelligent agents : papers from the November 8-10, Cambridge, Massachusetts.* Menlo Park, Calif., AAAI Press c1997.

Dautenhahn, K. and C. L. Nehaniv (1999). Living with Socially Intelligent Agents: A Cognitive Technology View. In. K. Dautenhahn. *Human Cognition and Social Agent Technology.* Amsterdam, John Benjamins.

Dautenhahn, K. and C. L. Nehaniv (2002). The Agent-Based Perspective on Imitation. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Dautenhahn, K. and C. Numaoka (1998). *Socially intelligent agents*.

Dautenhahn, K. and C. Numaoka (1999). Special issue: Socially intelligent agents, part II. *Applied Artificial Intelligence* **13**(3): 211-212.

Demiris, J. and G. Hayes (1997). Do Robots Ape? In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 28-30.

Demiris, J. and G. Hayes (2002). Imitation as a Dual-Route Process Featuring Predictive and Learning Components: A Biologically-Plausible Computational Model. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

den Boer, P. J. (1999). Natural selection or the non-survival of the non-fit. *Acta Biotheoretica* **47**: 83-97.

Dignum, F. and B. v. Linder (1997). Modelling Social Agents: Communication as Action. In. *Proceedings of the {ECAI}'96 Workshop on Agent Theories, Architectures, and Languages: Intelligent Agents {III}*. Berlin, Springer. **1193:** 205--218.

Doran, J. (2000). The Archeology of Artificial Societies. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep61.html

Dorigo, M., E. Bonabeau, et al. (2000). Ant algorithms and stigmergy. *Future Generation Computer Systems* **16**(8): 851-871.

Drogoul, S. e. b. P. a. A. (2000). *Robots as a Social Species: The MICRobES Project* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 139-141.

Dryer, D. C. (1997). Ghosts in the Machine: Personalities for Socially Adroit Software Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 31-36.

Durfee, E. H. (1991). *Organizations, Plans, and Schedules: {A}n Interdisciplinary Perspective on Coordinating {AI} Agents*. Journal of Intelligent Systems

Edmonds, B. (1997). Modeling Socially Intelligent Agents in Organisations. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 37-42.

Edmonds, B. (1998). Modeling socially intelligent agents. *Applied Artificial Intelligence* **12**(7-8): 677-699.

Edmonds, B. (1999). Capturing social embeddedness: A constructivist approach. *Adaptive Behavior* **7**(3-4): 323-347.

Edmonds, B. (1999). Gossip, Sexual Recombination and the El Farol Bar: modelling the emergence of heterogeneity. *Journal of Artificial Societies and Social Simulation* **2**(3).

Edmonds, B. (1999). Modelling Bounded Rationality In Agent-Based Simulations using the Evolution of Mental Models. In. T. Brenner. *Computational Techniques for Modelling Learning in Economics*, Kluwer**:** 305-332.

Edmonds, B. (2000). The Constructability of Artificial Intelligence (as defined by the Turing Test). *Journal of Logic Language and Information* **9**: 419-424.

Edmonds, B. (2000). *Developing Agents Who Can Relate to Us -- Putting Agents in our Loop via Situated Self-creation* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 55-60.

Edmonds, B. (2001). A bargaining model to simulate negotiations between water users - Commentary on: Sophie Thoyer, Sylvie Morardet, Patrick Rio, Leo Simon, Rachel Goodhue and Gordon Rausser (2001). *Jasss-the Journal of Artificial Societies and Social Simulation* **4**(2): U121-U125.

Edmonds, B. (2001). Learning appropriate contexts. In. *Modeling and Using Context, Proceedings*. **2116:** 143-155.

Edmonds, B. (2002). Developing Agents Who can Relate to Us: Putting Agents in Our Loop via Situated Self-Creation. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Edmonds, B. (2002). Reasoning about rational agents. *Jasss-the Journal of Artificial Societies and Social Simulation* **5**(1): U128-U130.

Edmonds, B. and K. Dautenhahn (1998). The Contribution of Society to the Construction of Individual Intelligence. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep42.html

Edmonds, B. and K. Dautenhahn (1999). Social intelligence. *Computational and mathematical organization theory* **5**(3): 199-202.

Edmonds, B. and K. Dautenhahn (2001). Special issue - Starting from Society: the Application of Social Analogies to Computational Systems. *Jasss-the Journal of Artificial Societies and Social Simulation* **4**(1): U3-U6.

Edmonds, B. and S. Moss (2001). The importance of representing cognitive processes in multi- agent models. In. *Artifical Neural Networks-Icann 2001, Proceedings*. **2130:** 759-766.

Edmonds, B. and S. Moss (2001). *The Importance of Representing Cognitive Processes in Multi-Agent Models*. Artificial Neural Networks - ICANN'2001, Vienna, Austria, Springer-Verlag, 759-766.

Elliott, C. and J. Brzezinski (1998). Autonomous agents as synthetic characters. *AI Magazine* **19**(2): 13-30.

Ephrati, E. and J. S. Rosenschein (1993). Multi-Agent Planning as a Dynamic Search for Social Consensus. In. *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI-93)*. {Chamb\'ery}, France**:** 423--429.

Falcone, R. and C. Castelfranchi (2001). The human in the loop of a delegated agent: The theory of adjustable social autonomy. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 406-418.

Faratin, P. (2002). Multi-Agent Contact Negotiation: Knowledge and Computation Complexities. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Fioretti, G. (2000). Recognition of investment opportunities and generation of investment cycles. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep62.html

Fitoussi, D. and M. Tennenholtz (2000). Choosing Social Laws for Multi-Agent Systems: Minimality and Simplicity. *Artificial Intelligence* **119**(1--2): 61--101.

Fitzpatrick, C. B. a. P. (2000). *That Certain Look: Social Amplification of Animate Vision* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 18-22.

Fone, L. N. (1999). Are We Having Fun Yet? Using Social Agents in Social Domains. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Frank, A., A. Stern, et al. (1997). Socially Intelligent Virtual Petz. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 43-45.

Fredslund, L. D. C. n. a. a. J. (2000). *How Does It Feel? Emotional Interaction with a Humanoid LEGO Robot* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 23-28.

French, R. M. (1990). Subcognition and the limits of the Turing Test. *Mind* **99**: 53-65.

Fritz, J. and K. Kotrschal (2002). On Avian Imitation: Cognitive and Ethological Perspectives. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Funk, P. and J. Lind (1997). What Is a Friendly Agent? In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 46-48.

Gasser, L. (1991). Social Conceptions of Knowledge and Action - Dai Foundations and Open Systems Semantics. *Artificial Intelligence* **47**(1-3): 107-138.

Gaussier, P., S. Moga, et al. (1997). From Perception-Action loops to Imitation Processes: A Bottom-up Approach of Learning by Imitation. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 49-54.

Gerber, C. (1997). An Artificial Agent Society Is More than a Collection of "Social" Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 55-57.

Gigerenzer, G. and D. G. Goldstein (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* **104**: 650-669.

Glos, J. W. (1999). Digital Augmentation of Keepsake Objects: A Place for Interaction of Memory, Story, and Self. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Gmytrasiewicz, C. L. L. a. P. J. (2000). *Decisions, Decisions ... and the Role of Emotions in the Process: A Formal Theory* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 96-101.

Goldman, C. V. and J. S. Rosenschein (1997). Multiagent Learning Systems and Expert Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 58-60.

Goodenough, O. R. (2002). Information Replication in Culture: Three Modes for the Transmission of Culture Elements through Observed Action. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Grand, S. (1999). The Emergence of Personality: How to Create Souls from Cells. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Granovetter, M. (1985). Economic-Action and Social-Structure - The Problem of Embeddedness. *American Journal Of Sociology* **91**(3): 481-510.

Gratch, J. (2000). *Socially Situated Planning* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 61-64.

Gratch, J. (2002). Socially Situated Planning. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Grohmann, A., R. Kopetzky, et al. (1999). Agents: An approach for dynamic process modelling. In. *Computing Anticipatory Systems*. **465:** 245-257.

Hall, L. E., L. Macaulay, et al. (1994). User Role in Problem Solving with Distributed Artificial Intelligent Systems. In. C. C. a. E. Werner. *Proceedings of the 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World : Artificial Social Systems*. Berlin, Springer. **830:** 324--337.

Hannoun, M., O. Boissier, et al. (2000). MOISE: An organizational model for multi-agent systems. In. *Advances in Artificial Intelligence*. **1952:** 156-165.

Heitsch, S., D. Hinck, et al. (2000). A New Look into Garbage Cans Petri Nets and Organisational Choice. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep63.html

Henrickson, L. (2000). Having a Sense of Ourselves: Technology and Personal Identity. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep64.html

Herman, L. M. (2002). Vocal, Social, and Self Imitation by Bottlenosed Dolphins. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Heyes, C. (2002). Transformational and Associative Theories of Imitation. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Heylighen, F. (1999). Collective intelligence and its implementation on the web: to develop a collective mental map

intelligence. *Computational and mathematical organization theory* **5**(3): 253-280.

Hofstadter, D. R. (1985). Analogies and Roles in Human and Machine Thinking. In. *Metamagical Themas*. New York, Basic Books.

Hogg, L. M. and N. R. Jennings (1997). Socially Rational Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 61-63.

Hogg, L. M. J. and N. R. Jennings (2001). Socially intelligent reasoning for autonomous agents. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 381-393.

Hudlicka, E. (2000). *Increasing SIA Architecture Realism by Modeling and Adapting to Affect and Personality* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 65-70.

Hudlicka, E. (2002). Increasing SIA Architecture Realism by Modeling and Adapting for Affect and Personality. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Hui, P. M., Y. R. Kwong, et al. (1999). Global behavior in a population of adaptive competitive agents. *Adaptive Behavior* **7**(3-4): 243-253.

Isbister, K. (2002). Party Hosts and Tour Guides: Using Nonverbal Social Cues in the Design of Interface Agents to Support Human-Human Social Interaction. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Jack, H. M. a. M. (2000). *Empirical Evaluation of Animated Agents in a Multi-Modal E-Retail Application*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 122-126.

Jacobsen, C. (1999). The process of crescive legitimation: Theory, simulation model, and three empirical tests. *Adaptive Behavior* **7**(3-4): 255-267.

Jacobson, R. (1998). Positioning the analysis: Memetics as a methodological tool to close the ranks between social and traditional history. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep43.html

Jarrold, W. (2000). *Towards the Essentials of Cyberlogical Testing: Methodologies for Inculcating Knowledge Based Systems with Social Intelligence* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 71-73.

Jebara, T., Y. Ivanov, et al. (2000). *Tracking Conversational Context for Machine Mediation of Human Discourse*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 74-76.

Jennings, N. R. (1999). Agent-based computing: Promise and perils. In. *Ijcai-99: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, Vols 1 & 2*: 1429-1436.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence* **117**(2): 277-296.

Johnson, W. L. (2000). *Socially Intelligent Agent Research at CARTE* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 77-82.

Jonker, C. M. and J. Treur (2001). Agent-based simulation of animal behaviour. *Applied Intelligence* **15**(2): 83-115.

Kalenka, S. and N. R. Jennings (1995). On Social Attitudes: {A} Preliminary Report. In. *International Workshop on Decentralised Intelligent and Multi-Agent Systems*. Krakow, Poland.

Kaminka, G. A. and M. Tambe (1997). Social Comparison for Failure Detection and Recovery in Multi-Agent Settings. In. *Proceedings of the 14th National Conference on Artificial Intelligence and 9th Innovative Applications of Artificial Intelligence Conference ({AAAI} 97/{IAAI}-97)*. Menlo Park, AAAI Press**:** 834--834.

Kaminka, G. A. and M. Tambe (1997). Towards Social Comparison for Failure Detection. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 64-66.

Kaminka, G. A. and M. Tambe (1998). What is wrong with us? Improving robustness through social diagnosis. In. *Fifteenth National Conference on Artificial Intelligence (Aaai- 98) and Tenth Conference on Innovative Applications of Artifical Intelligence (Iaai-98) - Proceedings***:** 97-104.

Kannapan, S. M. and K. M. Marshek (1992). A Schema for Negotiation between Intelligent Design Agents in Concurrent Engineering. *Ifip Transactions B-Applications in Technology* **4**: 1-20.

Kennedy, J. (1997). Minds and Cultures: Particle Swarm Implications. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 67-72.

Kennedy, J. (1999). Minds and cultures: Particle swarm implications for beings in sociocognitive space. *Adaptive Behavior* **7**(3-4): 269-287.

Kennedy, J. (2000). *Human and Computer Learning Together in the Exteriorized Particle Swarm* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 83-89.

Keyson, E. D. a. D. V. (2000). *Designing Collaboration in Consumer Products* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 52-54.

Kido, T. (1997). Is Social Intelligence Necessary for Agent-Agent Interaction? Toward Socially Intelligent Software Agents In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 73-74.

Kido, T. (1998). Grand Challenge Problems on Cross Cultural Communication --- Toward Socially Intelligent Agents. *Lecture Notes in Computer Science* **1435**: 108--??

Kido, T. (1998). Invited Contribution: Grand Challenge Problems on Cross Cultural Communication: Toward Socially Intelligent Agents. In. M. K. a. G. Wei{\ss}. *Proceedings of the 2nd International Workshop on Cooperative Information Agents {II}: Learning, Mobility and Electronic Commerce for Information Discovery on the Internet*. Berlin, Springer. **1435:** 108--112.

Kitano, H. and S. Tadokoro (2001). RoboCup Rescue - A grand challenge for multiagent and intelligent systems. *AI Magazine* **22**(1): 39-52.

Klawe, C. C. a. M. (2000). *Socially Intelligent Agents to Improve the Effectiveness of Educational Games* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 29-34.

Kozima, H. (2002). Infanoid: A Babybot that Explores the Social Environment. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Kubota, H. and T. Nishida (2000). *EgoChat Agent: A Talking Virtualized Member for Supporting Community Knowledge Creation*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 90-95.

Kubota, H. and T. Nishida. (2002). EgoChat Agent: A Talking Virtualized Agent that Supports Community Knowledge Creation. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Kummer, H., L. Daston, et al. (1997). The social intelligence hypothesis. In. W. e. al. *Human by Nature: between biology and the social sciences*. Hillsdale, NJ, Lawrence Erlbaum Associates**:** 157-17.

Lachman, R. (1998). AI, decision science, and psychological theory in decisions about people - A case study in jury selection. *AI Magazine* **19**(1): 111-129.

Landauer, C. and K. L. Bellman (1999). Computational Embodiment: Agents as Constructed Complex Systems. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Lanzola, G., L. Gatti, et al. (1999). A framework for building cooperative software agents in medical applications. *Artificial Intelligence in Medicine* **16**(3): 223-249.

Lee, C., M. Kearns, et al. (2000). Cobot in LambdaMOO: A Social Statistics Agent. In. *Proceedings of the 7th Conference on Artificial Intelligence ({AAAI}-00) and of the 12th Conference on Innovative Applications of Artificial Intelligence ({IAAI}-00)*. Menlo Park, CA, AAAI Press**:** 36--41.

Lee, R. M. (1997). A messenger model for navigating among bureaucratic requirements. In. *Thirtieth Hawaii International Conference on System Sciences, Vol 4 - Information Systems - Internet and the Digital Economy***:** 468-477.

Lenox, T., M. Lewis, et al. (2000). Task characteristics and intelligent aiding. In. *Smc 2000 Conference Proceedings: 2000 Ieee International Conference on Systems, Man & Cybernetics, Vol 1-5***:** 1123-1127.

Lieberman, H. (2002). Art Imitates Life: Programming by Example as an Imitation Game. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Lindsay, Marshall, et al. (2000). Modelling Agents using the Hotel Analogy: Sanitised for your Protection? Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep65.html

M. Mateas, A. S. (2002). Towards Integrating Plot and Character for Interactive Drama. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Machado, I. and A. Paiva. (2002). Me, My Character and the Others. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Mainzer, K. (2000). Society of knowledge - Interdisciplinary perspectives of information processing in virtual networks. In. *Classification and Information Processing at the Turn of the Millennium***:** 170-179.

Malsch, T. and I. Schulz-Schaeffer (1997). Generalized Media of Interaction and Interagent Coordination. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 75-80.

Marik, V., M. Pechoucek, et al. (2001). Social knowledge in multi-agent systems. In. *Multi-Agent Systems and Applications*. **2086:** 211-245.

Marik, V., O. Stepankova, et al. (2001). Acquaintance models for integration-oriented multi-agent systems. In. *Sccc 2001: Xxi International Conference of the Chilean Computer Science Society, Proceedings***:** 186-192.

Mark, G. and A. Voss (1997). Attractivity in Virtual Environments: Getting Personal with Your Agent. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 81-86.

Marková, I., Language, et al. (1978). *The social context of language*. Chichester ; New York, Wiley.

Marsella, S. C. (2000). *Pedagogical Soap* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 107-112.

Marsella, S. C. (2002). Pedagogical Soap: Socially Intelligent Agents for Interactive Drama. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Marsh, S. (1994). Trust in Distributed Artificial Intelligence. In. C. C. a. E. Werner. *Proceedings of the 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World : Artificial Social Systems*. Berlin, Springer. **830:** 94--114.

Mataric, M. J. (2002). Sensory-Motor Primitives as a Basis for Imitation: Linking Perception to Action and Biology to Robotics. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Mateas, M. (1997). Computational Subjectivity in Virtual World Avatars. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 87-92.

Mateas, M. and A. Stern (2000). *Towards Integrating Plot and Character for Interactive Drama*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 113-118.

McBreen, H. (2002). Embodied Conversational Agents in E-Commerce Applications. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

McBreen, H. M. and M. A. Jack (2001). Evaluating humanoid synthetic agents in e-retail applications. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 394-405.

Miceli, M. and C. Castelfranchi (1999). The Role of Evaluation in Cognition and Social Interaction. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Michaud, F. (2000). *Social Intelligence and Robotics*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 127.

Michaud, F. and C. Théberge-Turmel. (2002). Mobile Robotic Toys and Autism: Observations of Interaction. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Minsky, M. (2002). The Emotion Machine. http://web.media.mit.edu/~minsky

Mitchell, R. W. (2002). Imitation as a Perceptual Process. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Montemayor, J. and e. al. (2002). From Pets to Storyrooms: Constructive Storytelling Systems Designed with Children, for Children. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Moss, S. (1998). Social Simulation Models and Reality: Three Approaches. In. J. S. S. a. R. C. a. N. Gilbert. *Proceedings of the 1st International Workshop on Multi-Agent Systems and Agent-Based Simulation ({MABS}-98)*. Berlin, Springer. **1534:** 45--60.

Moss, S. (1999). Relevance, Realism and Rigour: A Third Way for Social and Economic Research. Manchester, Centre for Policy Modelling, Manchester Metropolitan University.

Moss, S. (2002). Challenges for Agent-Based Social Simulation of Multilateral Negotiation. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Moss, S. and B. Edmonds (1997). A knowledge-based model of context-dependent attribute preferences for fast moving consumer goods. *Omega-International Journal of Management Science* **25**(2): 155-169.

Moss, S. and B. Edmonds (1998). Modeling economic learning as modeling. *Cybernetics and Systems* **29**(3): 215-247.

Moss, S., C. Pahl Wostl, et al. (2001). Agent Based Integrated Assessment Modelling: The example of Climate Change. *Integrated Assessment* **2**(1).

Moulin, B. (1998). The social dimension of interactions in multi-agent systems. In. W. W. a. M. P. a. C. Zhang. *Agents and Multi-Agent Systems, Formalisms, Methodologies and Applications, Lecture Notes in Artificial Intelligence 1441*, Springer-Verlag: Heidelberg, Germany**:** 109--122.

Muhlenbrock, M., F. Tewissen, et al. (1997). A framework system for intelligent support in open distributed learning environments. In. *Artificial Intelligence in Education - Knowledge and Media in Learning Systems*. **39:** 191-198.

Nagao, K. and A. Takeuchi (1994). Social Interaction: Multimodal Conversation with Social Agents. In. *Proceedings of the 12th National Conference on Artificial Intelligence. Volume 1* Menlo Park, CA, USA, AAAI Press**:** 22--28.

Nehaniv, C. (2000). The Making of Meaning in Societies: Semiotic & Information-Theoretic Background to the Evolution of Communication. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep66.html

Nehaniv, C. L. (1997). What's the Story? Irreversibility, Algebra, Autobiographic Agents In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 150-156.

Nehaniv, C. L. and K. Dautenhahn (2002). The Correspondence Problem. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Nicolescu, M. N. and M. J. Mataric (2001). Learning and interacting in human-robot domains. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 419-430.

Noble, J. (1999). Cooperation, conflict and the evolution of communication. *Adaptive Behavior* **7**(3-4): 349-369.

Noble, J. and P. Todd (2002). Imitation or Something Simpler? Modelling Simple Mechanisms for Social Information Processing. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Now{\'{e}}, A., J. Parent, et al. (2001). Social Agents Playing a Periodical Policy. In. *Machine Learning: {ECML} 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5-7, 2001, Proceedings*, Springer. **2167:** 382--393.

Numaoka, C. (1997). Innate Sociability: Sympathetic Coupling. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 98-102.

Nwana, H. S. (1996). Software agents: An overview. *Knowledge Engineering Review* **11**(3): 205-244.

O' Riordan, C. and H. Sorensen (1999). An agent-based system for intelligent collaborative filtering. In. *Cooperative Information Agents Iii, Proceedings*. **1652:** 125-136.

Oliphant, M. (1999). The learning barrier: Moving front innate to learned systems of communication. *Adaptive Behavior* **7**(3-4): 371-383.

Oliphant, M. (2002). Rethinking the Language Bottleneck: Why Don't Animals Learn to Communicate? In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

O'Neill-Brown, P. (1997). Setting the Stage for the Culturally Adaptive Agent. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 93-97.

Onn, S. and M. Tennenholtz (1997). Determination of Social Laws for Multi-Agent Mobilization. *Artificial Intelligence* **95**(1): 155--167.

Ossowski, S. (1999). Co-ordination in artificial agent societies - Social structures and its implications for autonomous problem-solving agents - Introduction. In. *Co-Ordination in Artificial Agent Societies*. **1535:** 1-+.

Ossowski, S. and A. Garc{\'\i}a-Serrano (1999). Social Structure in Artificial Agent Societies: Implications for Autonomous Problem-Solving Agents. In. J. o. r. M. u. l. a. M. P. S. a. A. S.

Rao. *Proceedings of the 5th International Workshop on Intelligent Agents {V} : Agent Theories, Architectures, and Languages ({ATAL}-98)*. Berlin, Springer. **1555:** 133--148.

Ossowski, S. and A. GarciaSerrano (1996). A model of co-ordination support for unanticipated situations. In. *Advanced It Tools***:** 559-566.

O'Sullivan, D. and M. Haklay (2000). Agent-based models and individualism: is the world agent-based? *Environment and Planning A* **32**(8): 1409-1425.

Paiva, A., I. Machado, et al. (2001). The child behind the character. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 361-368.

Paiva, I. M. a. A. (2000). *The Child Behind the Character* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 102-106.

Palmer, R. G. e. a. (1994). Artificial Economic Life - A simple model of a stockmarket. *Physica D* **75**: 264-274.

Panzarasa, P., T. J. Norman, et al. (1999). Modelling Sociality in a BDI framework. In. *Proceedings of the 1st Asia-Pacific Conference on Intelligent Agent Technology*. Hong Kong.

Paolucci, M. and R. Conte (1999). Reproduction of nomative agents: A simulation study. *Adaptive Behavior* **7**(3-4): 307-321.

Paredes, A. L. and R. d. O. Martínez (1998). The Social Dimension of Economics and Multiagent Systems. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep44.html

Penny, S. (1997). Embodied Cultural Agents at the Intersection of Robotics, Cognitive Science and Interactive Art. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 103-105.

Penny, S. (1999). Agents as Artworks and Agent Design as Artistic Practice. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Pepperberg, I. M. (2002). Allospecific Referential Speech Acquisition in Grey Parrots (Psittacus erithacus): Evidence for Multiple Levels of Avian Vocal Imitation. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Perspective, A. B. and K. Dautenhahn (2000). Reverse Engineering of Societies. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep60.html

Persson, P. and e. al. (2002). Understanding Social Intelligence. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Persson, P., J. Laaksolahti, et al. (2000). *Anthropomorphism -- A Multi-Layered Phenomenon*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 131-135.

Persson, P., J. Laaksolahti, et al. (2001). Understanding socially intelligent agents - A multilayered phenomenon. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 349-360.

Petrushin, V. A. (2000). *Emotion Recognition Agents in Real World* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 136-138.

Petrushin, V. A. (2002). Creating Emotion Recognition Agents for Speech Signal. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Picault, S. (1998). A Multi-Agent Simulation of Primate Social Concepts. In. H. Prade. *Proceedings of the 13th European Conference on Artificial Intelligence ({ECAI}-98)*. Chichester, John Wiley & Sons**:** 327--330.

Pistolesi, G. and P. Modesti (2001). Complexity and state-transitions in social dependence networks. *Discrete Dynamics in Nature and Society* **6**(1): 49-+.

Pizzutilo, S. and e. al. (2002). Cooperative Interface Agents. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Pizzutilo, S., B. D. Carolis, et al. (2000). *Cooperating with Personality-Rich Interface Agents*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 142-147.

Prendinger, H. and M. Ishizuka (2001). Let's talk! Socially intelligent agents for language conversation training. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 465-471.

Prendinger, H. and M. Ishizuka (2001). Simulating affective communication with animated agents. In. *Human-Computer Interaction - Interact'01***:** 182-189.

Price, B. and C. Boutilier (2000). Imitation and Reinformcement Learning with Heterogeneous Actions. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep67.html

Prietula, M. and K. Carley (1997). Agents, Trust, and Organizational Behavior. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 146-149.

Prietula, M. J. and K. M. Carley (2000). *Boundedly Rational and Emotional Agents: Simulating DAISY and DAIDO Trust Models*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 148-149.

Pynadath, D. V., M. Tambe, et al. (2000). *Electric Elves: Immersing an Agent Organization in a Human Organization*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 150-154.

Pynadath, D. V. and M. Tambe. (2002). Electric Elves: Adjustable Autonomy in Real-World Multi-Agent Environments. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Rahmer, J. and A. Voxx (1997). Brownies as Assistants. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 106-108.

Rampoldi-Hnilo, L. A. (2000). *Monitoring Our Online Representations -- Avatars, the Digital Mirror* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 155-160.

Rizzo, P., M. Veloso, et al. (1997). Personality-Driven Social Behaviors in Believable Agents. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 109-114.

Robins, J. (2000). *The Inquiry Page: A K-12 Collaboratory Using Socially Intelligent Agents* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 161-163.

Rodríguez-Aguilar, J. A. and C. Sierra (2002). Enabling Open Agent Institutions. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Rodriguezrosello, L. (1993). Research-and-Development on Advanced Learning Technologies in Europe. In. *Computer Mediated Education of Information Technology Professionals and Advanced End-Users*. **35:** 3-14.

Rolls, E. T. (2000). Precis of the brain and emotion. *Behavioural and Brain Sciences* **23**(2): 177-.

Rouchier, J. (2002). Social Intelligence for Computers: Making Artificial Entities Creative in their Interactions. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Salvador, L.-L. (1997). Assimilation, Imitation and the Elementary Social Fact. Towards a Definition of Social Interactions. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 115-117.

Sammut, C., S. Hurst, et al. (2002). Learning to Fly. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Sánchez-Segura, M.-I., R. Imbert, et al. (1998). Modelling and Evolution of Social Trends in Virtual Environments. Workshop on Socially Situated Intelligence at SAB'98, Zurich, http://cfpm.org/cpmrep45.html

Scassellati, B. (2000). *Theory of Mind... for a Robot* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 164-168.

Scerri, P., M. Tambe, et al. (2000). *DON'T CANCEL MY BARCELONA TRIP: Adjusting Autonomy of Agent Proxies in Human Organizations*. Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 169-173.

Scheeff, M. and e. al. (2002). Experiences with Sparky, a Social Robot. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Schillo, M., S. Allen, et al. (2000). Socially Competent Business Agents with Attitude Using Habitus-Field Theory to Design Agents with Social Competence. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep68.html

Schillo, M., J. u. r. Lind, et al. (1999). The Social Interaction Framework \\ System Description and User's Guide to a Multi-Agent System Testbed

Research Report RR-99-02. Deutsches Forschungszentrum f{\"u}r K{\"u}nstliche Intelligenz GmbH\\Erwin-Schr{\"o}dinger Stra{ss}e\\Postfach 2080\\67608 Kaiserslautern\\Germany, Deutsches Forschungszentrum f{\"u}r K{\"u}nstliche Intelligenz**:** 30.

Seco, J. C., C. Pinto-Ferreira, et al. (1998). A society of agents in environmental monitoring. In. *From Animals to Animats 5***:** 447-452.

Sekiyama, K. and T. Fukuda (1997). Toward Social Robotics. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 118-124.

Sengers, P. (1997). Towards Socially Intelligent Agent-Building. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 125-130.

Sengers, P. (1999). Narrative Intelligence. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Shmeil, M. A. H. and E. Oliveira (1997). The establishment of partnerships to create virtual organizations: A multiagent approach. In. *Re-Engineering for Sustainable Industrial Production***:** 284-294.

Shoham, Y. and M. Tennenholtz (1992). On the Synthesis of Useful Social Laws for Artificial Agent Societies. In. P. R. a. P. Szolovits. *Proceedings of the Tenth National Conference on pages = {276--281}, Artificial Intelligence*. Menlo Park, California, AAAI Press

American Association for Artificial Intelligence.

Shoham, Y. and M. Tennenholtz (1993). *On Social Laws for Artificial Agent Societies: Off-Line Design*. Artificial Intelligence, to appear.

Simon, H. (1986). The failure of armchair economics (Interview). *Challenge* **29**(5): 18-25.

Slator, B. M. and G. Farooque (1998). The agents in an agent-based economic simulation model. In. *International Society for Computers and Their Applications 11th International Conference on Computer Applications in Industry and Engineering***:** 175-178.

Sloman, A. (1999). Architectural Requirements for Human-like Agents Both Natural and Artificial. (What sorts of machines can love?). In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Sloman, A. (2002). How many separately evolved emotional beasties live within us?" In. R. Trappl, P. Petta and P. S. *Emotions in Humans and Artifacts*. Cambridge, MA, MIT Press.

Staller, A. and P. Petta (2000). Introducing Emotions into the Computational Study of Social Norms. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep69.html

Tadokoro, S., H. Kitano, et al. (2000). The RoboCup-Rescue: An international cooperative research project of robotics and AI for the disaster mitigation problem. In. *Unmanned Ground Vehicle Technology Ii*. **4024:** 303-312.

Takadama, K., T. Terano, et al. (1999). Making organizational learning operational: implications from learning classifier systems. *Computational and mathematical organization theory* **5**(3): 229-252.

Tambe, M. (1997). Towards Flexible Teamwork In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 131-133.

Tambe, M., J. Adibi, et al. (1999). Building agent teams using an explicit teamwork model and learning. *Artificial Intelligence* **110**(2): 215-239.

Tenhaaf, N. (2000). *Where Surfaces Meet* Socially Intelligent Agents: The Human in the Loop, AAAI Fall Symposium, North Falmouth, MA, AAAI Press, 174-179.

Tenhaaf, N. (2002). Perceptions of Self in Art and Intelligent Agents. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.

Tennenholtz, M. (1999). On social constraints for rational agents. *Computational Intelligence* **15**(4): 367-383.

Todd, P. M. and B. Borges (1997). Designing Socially Intelligent Agents for the Ultimatum Game. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 134-136.

Traum, D. R. (1995). Social Agency and Rationality. In. M. Fehling. *Working Papers of the {AAAI} Fall Symposium on Rational Agency: Concepts, Theories, Models, and Applications* Menlo Park, California, American Association for Artificial Intelligence**:** 150--153.

Vainio, M., A. Halme, et al. (1997). Robot Societies --From Simulator Studies towards a Real World Application. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 137-139.

Visalberghi, E. and D. Fragaszy (2002). Do Monkeys Ape? - Ten Years After. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Vogt, S. (2002). Dimensions of Imitative Perception-Action Mediation. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Vorn, B. (1999). Machine-Mediated Communication: Agents of Representation. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*. Amsterdam, John Benjamins.

Watt, S. N. K. (1996). Artificial societies and psychological agents. *Bt Technology Journal* **14**(4): 89-97.

Weber, R., D. W. Aha, et al. (2001). Intelligent lessons learned systems. *Expert Systems with Applications* **20**(1): 17-34.

Wellens, A. R. and M. D. McNeese (1999). The social psychology of intelligent machines: a research agenda revisited. In. *Proceedings of the Eighth International Conference on Human-Computer Interaction*. **1:** 696--700.

Werner, E. (1989). Cooperating Agents: {A} Unified Theory of Communication and Social Structure. In. L. G. a. M. N. Huhns. *Distributed Artificial Intelligence Vol. II*. London, Pitman**:** 3--36.

Werner, E. (1999). The Ontogeny of the Social Self. Towards a Formal Computational Theory. In. K. Dautenhahn. *Human Cognition and Social Agent Technology*, John Benjam**:** 263-300.

Werry, I., K. Dautenhahn, et al. (2001). Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy. In. *Cognitive Technology: Instruments of Mind, Proceedings*. **2117:** 57-74.

Whiten, A. (2002). Imitation of Sequential and Hierarchical Structure in Action: Experimental Studies with Children and Chimpanzees. In. K. Dautenhahn and C. L. Nehaniv. *Imitation in Animals and Artifacts*. Cambridge, Mass, MIT Press.

Wilkes, D. M., R. T. Pack, et al. (1997). HuDL, A Design Philosophy for Socially Intelligent Service Robots. In. K. Dautenhahn. *Socially Intelligent Agents, AAAI Fall Symposium*. Cambridge, MA, AAAI Press. **FS-97-02:** 140-145.

Wilson, M. and M. Neal (2001). Diminishing returns of engineering effort in telerobotic systems. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 459-465.

Wimsatt, W. (1972). Complexity and Organisation. In. K. Schaffner and R. Cohen. *Studies in the Philosophy of Sciences*. Dordretch, Reidel. **20:** 67-86.

Wooldridge, M. (2000). *Computationally Grounded theories of Agency*. 4th International Conference on MultiAgent Systems – ICMAS-2000, Los Alamitos, CA, IEEE Computer Society, 13-20.

Wooldridge, M. and N. R. Jennings (1998). *Pitfalls of Agent-Oriented Development*. Second International Conference on Autonomous Agents, ACM Press

Wright, I. (2000). The Society of Mind Requires an Economy of Mind. Starting from Society: the application of social analogies to computational systems, AISB symposium, Birmingham, http://cfpm.org/cpmrep70.html

Yamaguchi, T. (2000). Interactive self-reflection based reinforcement learning for multiagent coordination. In. *Intelligent Autonomous Systems 6***:** 561-566.

Ye, M. and K. M. Carley (1995). Radar-Soar: Towards and artificial organization composed of intelligent agents. *Journal of Mathematical Sociology* **20**(2-3): 219-246.

Ye, Y. M., S. Boies, et al. (2001). Agents-supported adaptive group awareness: Smart distance and WWWaware. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans* **31**(5): 369-380.

Young, R. M. (2002). The Cooperative Contract in Interactive Entertainment. In. K. Dautenhahn, A. H. Bond, L. Cañamero and B. Edmonds. *Socially Intelligent Agents - Creating Relationships with Computers and Robots*. Boston, MA, Kluwer Academic. **3**.