

Narrative Intelligence from the Bottom Up: A Computational Framework for the Study of Story-Telling in Autonomous Agents

Kerstin Dautenhahn & Steven J. Coles

Adaptive Systems Research Group, Department of Computer Science, University of
Hertfordshire, College Lane, Hatfield Herts AL10 9AB, UK

K.Dautenhahn@herts.ac.uk

<http://homepages.feis.herts.ac.uk/~comqkd>

Abstract

This paper addresses *Narrative Intelligence* from a bottom up, Artificial Life perspective. First, different levels of narrative intelligence are discussed in the context of human and robotic story-tellers. Then, we introduce a computational framework which is based on minimal definitions of stories, story-telling and autobiographic agents. An experimental test-bed is described which is applied to the study of story-telling, using robotic agents as examples of situated, autonomous *minimal agents*. Experimental data is provided which support the working hypothesis that story-telling can be advantageous, i.e. increases the survival of an autonomous, autobiographic, minimal agent. We conclude this paper by discussing implications of this approach for story-telling in humans and artifacts.

1 Introduction and Background

The issue of *Narrative Intelligence* has recently attracted a lot of attention in the Artificial Intelligence (AI) community (see e.g. Wyer 1995; Mateas & Sengers 1999; Sengers 2000; Mateas & Sengers, forthcoming) and resulted in new interfaces for human-computer interaction which are either supporting human narrative intelligence (e.g. Umaschi-Bers & Cassell 2000), or which result in software agents that can (autonomously or in interaction with a user) enact a story (Aylett 1999), e.g. in interactive learning environments (Machado et al. 1999). Literature in narrative psychology (Bruner 1987, 1990, 1991) and developmental psychology (Nelson 1986, 1993) stress the central role of stories in the development of a social self in human beings, see Bruner's list of characteristics of stories which are different from other types of memories (Bruner 1991), and their implications for AI systems (Sengers 2000). Stories are efficient vehicles for communicating about others and oneself, some even define the human self as a "centre of narrative gravity" (Dennett 1989).

Often, research on stories and narratives is strongly biased towards *human* narrative intelligence, addressing spoken or written language, and the role of narrative in human thought, e.g. (Turner 1996). Research by Dautenhahn and Nehaniv (Dautenhahn & Nehaniv 1998; Nehaniv & Dautenhahn 1997; Nehaniv & Dautenhahn 1998; Dautenhahn 1998; Dautenhahn 1999) takes a radically different approach: instead of starting from sophisticated (language-based) forms of human narrative we are interested in the "evolutionary" origin of stories. This raises questions e.g. about the precursors of "stories" in non-human primates and other animals (discussed in Dautenhahn 1999b, Dautenhahn forthcoming), or how one can interpret "stories" from an Artificial Life, agent-perspective.

This article presents a computational framework for the investigation of story-telling for autonomous agents. Previous work (Coles & Dautenhahn 2000) investigated techniques of memory organisation and representation of stories. In this paper we are defining and experimentally exploring *minimal* definitions of stories, and minimal conditions which indicate how stories might provide an "evolutionary" advantage for story-telling agents in contrast to their non-story-telling "relatives".

The work presented in this article is on the one hand related to research in robotics that is developing memory-based controllers for autonomous robots. In this area, different approaches are investigated, ranging from behaviour-based local representations, e.g. (Kuipers 1987), (Mataric 1992), (Michaud & Mataric 1998; 1999), to dynamical systems approaches typically involving recurrent neural networks, e.g. (Tani 1996), (Billard & Hayes 1999). On the other hand, our research is related to work in psychology and cognitive science studying episodic and autobiographic memory in humans, e.g. (Tulving 1983; 1993), (Nelson 1986), (Conway 1996), (Glenberg 1997).

We like to point out, that the goal of this article is not to present a unified theory of episodic memory integrating findings from robotics and psychology. Instead, our target area is the emerging field of *Narrative Intelligence*, using the method of agent-based modelling (more specifically, using an embodied simulated robot), following an experimentally driven methodology for behavioural design that is well established in behaviour-based robotics and Artificial Life, cf. (Langton 1995), (Arkin 1998). Generally, this approach is grounded in a bottom-up approach towards understanding animal and human intelligence (Brooks 1999; Pfeifer & Scheier 1999). Similarly, agent-based simulation models e.g. used by Epstein and Axtell (1996), and many other contributions that can be found in the area of artificial societies and social simulation, attempt to understand phenomena observed in human societies by using a level of description that abstracts from the complexity of morphology and behaviour of single human agents and studies emergent effects of interactions among *simple* agents instead.

It is also important to note that although we study a single agent scenario, this work is part of an ongoing project that investigates communication and story-telling in dyads or groups of agents. Our approach is not meant to oppose the view of human intelligence and human cognition as being inherently social, as it has been argued e.g. in (Vygotsky 1986), (Wertsch 1985), (Brothers 1990), (Brothers 1997). On the contrary, we proposed previously an approach towards social robotics based on the notion that intelligence is socially embedded (Dautenhahn 1995; Dautenhahn 1999; Dautenhahn & Billard 1999). However, in addition to acknowledging the important role of the social environment in human development, we assume that the *social brain* of an autonomous agent (natural or artificial) depends to some extent as well on a variety of either innate ("hard-wired"), or individually learnt skills or developed capacities (cf. language acquisition). In the context of the work presented in this article we assume that it is plausible to discuss the basic mechanism of remembering a sequence of previous actions with respect to a *single* agent.

2 Levels of Narrative Intelligence

How can we discuss narrative intelligence below the level of human narrative intelligence? In Dautenhahn (1998) we introduced different levels of narrative complexity, starting from the most simple (trivial) case of a narrative agents, and moving towards high-level forms of human narrative intelligence. In the following we provide a brief summary, outlining the implications for robotic story-tellers.

Type 0: In Wyer (1995) Schank and Abelson discuss the *grandfather model* of memory, describing an agent that is telling the same story over and over again. An agent which is always telling/replying a single story, no matter what story it is told, would not be considered to be intelligent since it does not react or adapt to another agent. Building a robot so that it always responds with exactly the same (or a slightly modified) story, triggered by a particular event in the environment (e.g. input by a human user) is quite straightforward. Such a system would however more resemble a doorbell than a communicative robot. A narrative agent of Type 0 is an extreme form of a Type I agent.

A Type I agent can tell a great variety of stories, but the stories are not situated in the conversational context, i.e. the agent *randomly* selects a single story from its story-base and tells it in exactly the same way as it was stored. A Type I narrative robot would not be perceived as very communicative by human interaction partners. Users are likely to appreciate the responsiveness of the robot, but the lack of similarity between what the human communicates and how the robot responds does not support a satisfactory interaction/dialogue.

Type II narrative agent: It selects a story which fits best the current context, i.e. the story which is the best "response" to a story. A famous, early AI program ELIZA (Weizenbaum 1965) can be regarded as a dialogue agent with the same underlying principle. Thus, Type II narrative agents like ELIZA (if they possess a huge story-base or appropriate story-generation mechanisms) are able to produce responses

which can be quite complex, and with a low repetition rate. Such agents can be considered believable, but they have a pre-specified response/story-base, and, most important of all, they do not "listen" (understand). ELIZA does not understand and incorporate the user's stories in its own story-base in a meaningful way, i.e. the user's stories are not integrated in ELIZA's individual story-base. Type II agents are used successfully in many user-interface systems supporting human narrative intelligence (e.g. Umaschi-Bers and Cassell 2000). As long as such a level of interaction is sufficient, mobile robots can employ such systems, in particular in application areas with a restricted discourse domain.

Type III narrative agent: it tells and listens to stories, "understands" them in the sense that it is able to interpret the meaning and content of the story and retrieves from its own story-base the most similar story which is then adapted in order to produce an appropriate response. Script-like encoding and representation of experiences and case-based reasoning-like mechanisms for indexing, finding and adapting stories are well known AI techniques which can serve such a purpose, see also a discussion of story-telling by computer in Sharples (1997). Type III narrative agents can potentially make interesting and believable story-tellers.

Type IV narrative agent: This term refers to narrative intelligence as humans and possibly other animals possess them (Dautenhahn 1999b, Dautenhahn forthcoming), the most advanced form of narrative agents. What makes them special is that their story-telling ability is linked to a *historical*, autonomous agent, and its autobiography. Animal stories are either represented implicitly in their body shapes and behaviours which are reflecting their -phylogenetic (evolutionary) and ontogenetic (developmental) - history, or, of particular interest in the context of this paper, explicitly by communication, as humans and possibly other non-human animals do. Stories told by humans are rich and of such a complexity and variability that it is difficult to define human narrativity (but see Bruner 1991), and to separate it from the living body which such narratives emerge from. However, embedding such a story-telling capacity in virtual agents like Creatures (Grand et al. 1997), which have a simulated life-time, a rich behaviour repertoire and an artificial biology, might one day lead to artificial agents which can tell us stories (almost) as rich and interesting as those told by humans.¹ Stories can also be expressed non-verbally, using gestures or employing other sensor modalities (like bats communicating by ultrasound, insects using chemical or infrared signals etc.). Thus, story-telling autobiographic agents can provide rich and meaningful interaction, empowering humans by providing complex means to create stories and communicate meaning.

3 Definitions

This section provides core definitions necessary for our computational framework.

3.1 Autobiographic memory

Def.: An agent possesses an autobiographic memory if it can create and access information about sequences of actions which it experienced during its lifetime.

The complexity of such a memory can range from a simple linear sequence of actions/perceptions to highly structured and semantically organised systems as we know them from human autobiographic memory. In (Dautenhahn & Nehaniv 1998) we discuss different classes of information which can be stored in an autobiographical memory, e.g. a total "data log" of all perceptions and actions, the selection of only "meaningful" actions etc. Note, that our minimal definition of autobiographic memory does not include any notion of meaning or relevance of remembered events. Thus, we do not require that what the agent remembers is useful in the current context, nor that the agent has any notion of meaning or relevance of current or past situations at all. Also, autobiographic memory can be implicit in the body shape and behaviour of an agent, representing its phylogenetic and ontogenetic history of an agent, e.g. of a single-cell organism like *Paramecium*.

3.2 Autobiographic agents

¹ See a formal algebraic framework for story-telling artificial life agents in Nehaniv & Dautenhahn (1997, 1998).

Def.: Autobiographic agents are agents which are embodied and situated in a particular environment (including other agents), and which dynamically reconstruct their individual *history* (autobiography) during their lifetimes (Dautenhahn 1996).

In the New/Nouvelle Artificial Intelligence community ((Brooks 1999), (Pfeifer & Scheier 1999)) the term *situated* is generally defined as an agent that is surrounded by the real world and operates based on sensory perception that it gains in interaction with the environment, rather than operating upon abstract representations of reality (Arkin 1998:26). The term *embodiment* is often used to simply describe the fact that an agent is realised as a physical robot or an agent (Pfeifer & Scheier 1999:649). However, the issue of how embodiment can be defined is controversial (see discussions in Dautenhahn, 1999). In this paper we use the term embodiment in a more quantitative sense, referring to the definition developed by Tom Quick, Kerstin Dautenhahn and Chrystopher Nehaniv (Quick et al. 1999). Here, embodiment is based on channels of mutual perturbation, i.e. an agent is embodied in a particular environment if the environment can perturb the agent and if the agent can perturb the environment. Software agents, such as the simulated agent which we study, can therefore be embodied, although their degree of embodiment is lower than the degree of embodiment of physical objects, i.e. robots and biological organisms. As a special (trivial) case of an embodied system we might even consider a rock, since it influences the environment and in return is influenced by the environment. However, although the rock's shape does reflect its *history*, it is not an autobiographic agent since it cannot go back in time, namely reconstruct its history. Inanimate physical objects do have a temporal horizon (a past, a present and a future, see Nehaniv et al. 1999), but they cannot manipulate it, i.e. they cannot remember anything (let al.one making predictions of the future). Similarly, the body of animals and even plants reflects their phylogenetic and ontogenetic history. For many animal species we clearly know that they can remember the past and/or predict/manipulate the future (humans and other primates are good examples, possibly including elephants, dolphins, whales, parrots and other non-human animals). For artificial systems it is easier than for biological systems to draw the line between *reactive* systems (living exclusively in the here and now), and *post-reactive*, temporally grounded systems (Nehaniv et al. 1999). The computational framework presented in this paper and experiments on synthesising autobiographic agents might clarify the boundaries.

Note, that our minimal definition of autobiographic agents does not require that the agent adapts a remembered (previous) story to the current situation. Looking at (from what we know) one of the most complex forms of autobiographic memory, namely human autobiographic memory, we know that *reconstruction* has a very complex form: for example in human dialogue humans communicate and understand in stories, relating current stories to similar remembered ones, adapting old stories to new situations (Schank & Abelson 1977, Schank & Abelson in: Wyer 1995), creating new stories and adding them to their memory. In our minimal definition we do not assume any reconstruction of this kind. Reconstruction in its simplest version means the ability to maintain, expand and access memories of previous situations. The minimal definition of autobiographic memory does however imply that the agent is able to relate current to previous experiences by finding *similar* situations. However, no assumptions are made how complex this similarity measure needs to be.

3.3 Stories

Since we are interested in minimal definitions, we do not distinguish specifically between stories and narratives. We define stories as follows:

Def.: Stories are sequences of actions, expressed by an autonomous agent (including movements as well as "speech acts"), which can be related to previous situations in the agent's autobiographical memory.

3.4 Story-telling

Def.: An agent is said to "tell a story" if current perceptions (in conjunction with internal states of the agents) trigger the agent to retrieve and express a story from its autobiography.

Note that story-telling need not be triggered by another agent, it can be triggered by the agent himself. For example the agent can encounter a situation which reminds it of a similar previous situation. In this way the agent can tell a story to itself, rather than to other agents (Nehaniv 1997). Thus, in the minimal definition we do not assume that the agent can distinguish between itself and other agents. Other agents are simply part of the environment. The experimental work described in this paper is based on experiments with one agent.

4 Working Hypothesis

We use an experimental framework in order to study different hypotheses important to a bottom-up approach towards narrative intelligence (cf. Epstein and Axtell's bottom up approach to social sciences, using computer simulations for the study of social and cultural phenomena.). In this paper we investigate the hypothesis that story-telling can provide an advantage to an autonomous, autobiographic *minimal agent*. With minimal agent we mean an autobiographic agent with minimal knowledge of the environment, minimal mechanisms of reconstructing its autobiography, without any notion of "meaning" or "usefulness" of the stories it remembers. Moreover, remembering and story-telling is not triggered by another agent, it is triggered at regular intervals by the agent's own actions. Thus, a minimal agent is substantially different from the (human) subjects in narrative psychology. Although, as discussed in (Dautenhahn 1999b; Dautenhahn forthcoming), we strongly believe that social conditions were the major driving forces for the evolution of story-telling in human evolution (stories as the most efficient way to communicate with others and manage increasingly complex social dynamics), *basic* mechanisms of memory-organisation and story-telling are likely to have occurred first in an individual. As an analogy, from an evolutionary perspective, this relates to conditions when the first animal might have had its first experience of story-telling, possibly without any other agents in the vicinity that possessed the same story-telling capabilities. Thus, we assume that the origin of story-telling mechanisms must have provided an advantage to the individual, rather than the social group (the latter certainly accelerated the evolution of story-telling, once a population of story-telling agents, and culture, were established). Note, that our minimal agent does not know the meaning of the remembered story, nor how to relate this previous story to the current situation. Under such minimal conditions, can story-telling provide an advantage so that it enhances the survival of the individual? If yes, what exactly are these advantages?

5 Experimental Test-bed

A preliminary version of the experimental test-bed was first described in (Coles & Dautenhahn 2000). The experiments are conducted using the WEBOTS (*Cyberbotics*²) simulator which is widely used in autonomous agents research. The results reported in this paper are only based on the simulation experiments, but one important reason for choosing WEBOTS was that results can be relatively easily tested with physical KHEPERA robots (developed by *K-Team*³), and future experiments will exploit this option. The simulator was designed to match the realistic conditions of real-world robotic experiments as much as possible. To give an example: experimental runs with real robots in real worlds are never 100 percent reproducible, due to a range of physical properties of the agent itself and the environment (noise, friction, temperature sensitivity of sensors etc.). In order to account for these real world conditions, WEBOTS adds 10 percent of white noise to measurements of light and proximity. As a result of this, each simulation run is different from the other, even when identical experimental parameters and initial conditions are used. Thus, when changing controllers we can only compare runs under the same general experimental conditions (starting point, properties of environment etc.), the runs themselves are never *exactly* the same even if we keep the control program fixed. Figure 1 shows a simulated WEBOTS robot.

² <http://www.cyberbotics.com/>

³ <http://www.k-team.com/>

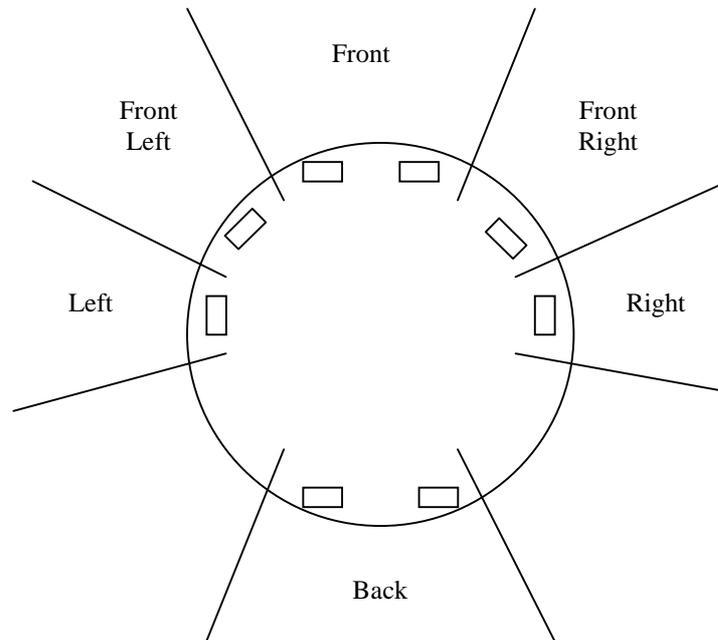


Figure 1: Schematic drawing of the WEBOTS (Khepera) robot used in the experiments. The robot possesses eight infra-red and eight light sensors and two wheels/motors (not shown). In-built odometry sensors measure the speed of the wheels. The infra-red sensors measure distance from walls/obstacles. For the purpose of our experiments we grouped the sensors so that they provide information about obstacles on the Left, Front Left, Front, Front Right, Right and Back of the agent.

The robot's behaviour-based control architecture (cf. subsumption architecture) is shown in Figure 2. Figure 3 shows the overall control architecture. Two light sources are located in the environment (see Figure 5). The robot explores the environment (simple Braitenberg-style obstacle avoidance, Braitenberg 1984) until its energy level drops below a certain threshold, it then switches to wall-following behaviour. The control program was designed so that the robot could survive for a reasonable lifetime in environment A shown in Figure 5, i.e. exploring the environment and recharging several times before it is running out of energy and "dies". Thus, the control program was designed so that the robot was well adapted to its environment, good enough to go through several cycles of exploration and homing. Parameters of the environment and control program were adjusted so that the robot was not "immortal", so that the environment still provided a "challenging" context. Once the robot runs out of energy a trial is ended. Trajectories and lifetime at the moment of "death" are recorded. The robot is tested in two different modes: using the control architecture (purely *reactive control*) as shown in Figure 3 (top part), or using the same control program extended by memory and story-telling mechanisms (*post-reactive control*, complete architecture shown in Figure 3).

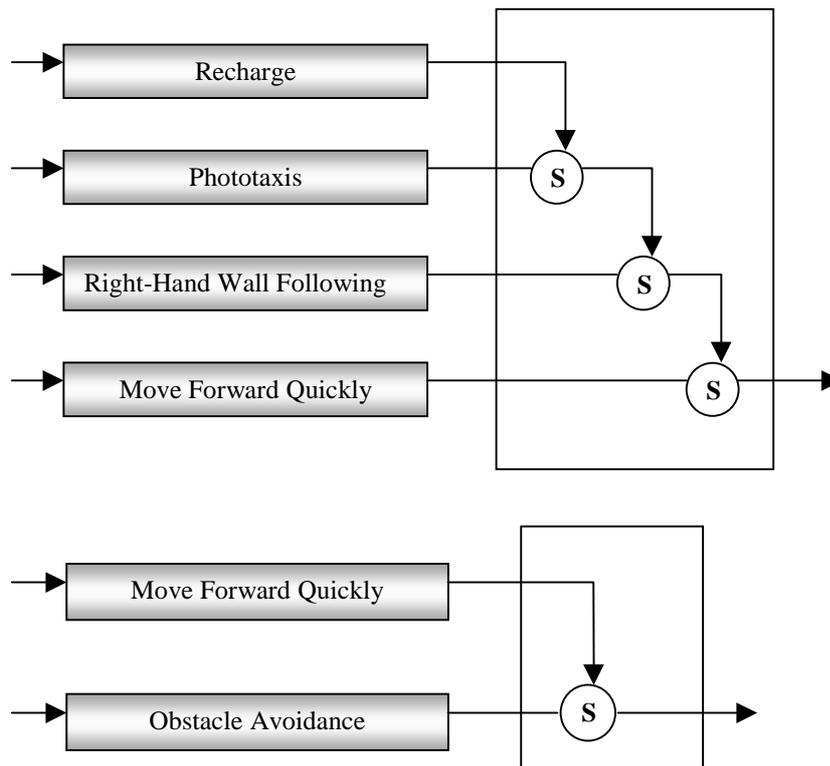


Figure 2: Behaviour hierarchies for homing (top) and exploration (bottom) mode. A subsumption type architecture is used (Brooks 1986). Higher-level behaviours inhibit or override lower-level behaviours. Obstacle avoidance is based on a simple Braitenberg algorithm.

The memory of the robot is organised in situation-action-situation triples $\langle S, A, S' \rangle$, shown in figure 4, inspired by discussions on mental maps (Kuipers 1982), and in particular the *Tour* model using a spatial semantic hierarchy (Kuipers & Levitt 1988) as a model of human path planning and navigation. Similar ideas are explored in robot navigation architectures, e.g. recognition-triggered response memory (Franz & Mallot 2000). The one-dimensional, circular memory of the robot (length 9000) stores the situation the robot is in at a particular simulation step, the action it is taking, and the new situation the robot finds itself in after taking that action. The complete autobiographic memory of the agent can at any given moment in time be saved in a file (Coles & Dautenhahn 2000), e.g. for later analysis. The history of a robot is constructed taking into account either a situation or an action (identified by an integer number), and the speeds of each of the wheels. Situation numbers are defined in terms of both light and proximity measures in order to represent situations such as very near, near or far. Action numbers are defined and memorised in terms of wheel speed. The robot's autobiography is in this way dynamically constructed and updated.

The robot writes its current action in a *memory buffer*. At a fixed interval (every 300 simulation steps, unless the robot is already in a state of remembering action sequences) the robot reads this memory buffer and searches its memory for a previous entry of the same action. The first encounter of the action (starting from the beginning of the memory array) is then taken as the starting point for the reconstruction of the action sequence. This stage can be labelled *day-dreaming* since the robot does not use its sensors, it is "blindly" repeating the sequence of actions (rehearsing the sequence of wheel speeds), triggered by the action in the memory buffer. Precaution was taken so that if the robot collides with an obstacle during daydreaming, it "wakes up" and continues with homing or exploration behaviour. Otherwise, if undisturbed, the robot wakes up automatically after 400 (remembered) action steps. The energy level is decreased for each step of the robot. If the robot is in homing mode and it is close to a light source then it

stops and recharges until its initial energy level is fully restored. Figure 6 shows an example of a robot remembering a previous sequence of actions. Since action numbers are defined and memorised in terms of wheel speed, the recalled trajectory is identical to the original one.

Although not necessary in our one-agent scenario, we implemented the memory buffer since the same buffer can also be read and modified by other agents, e.g. the control architecture can be used for a single agent which reminds itself of an action, as well as two or more agents using the memory buffer in a “dialogue”, i.e. mutually triggering responses based on entries in the memory buffer. The robot was tested in seven different runs for each experiment, starting in the left lower corner of the environment (see Fig. 5). In one series (experiment A1) the robot only used the reactive part of its control architecture, in the other series (experiment A2) its story-telling mechanisms. We then repeated the experiments in environment B, shown in figure 5, which lead to experiments B1 and B2. Lifetimes at the moment of “death” were tested statistically (standard t-Test), and trajectory patterns were compared qualitatively (figure 7).

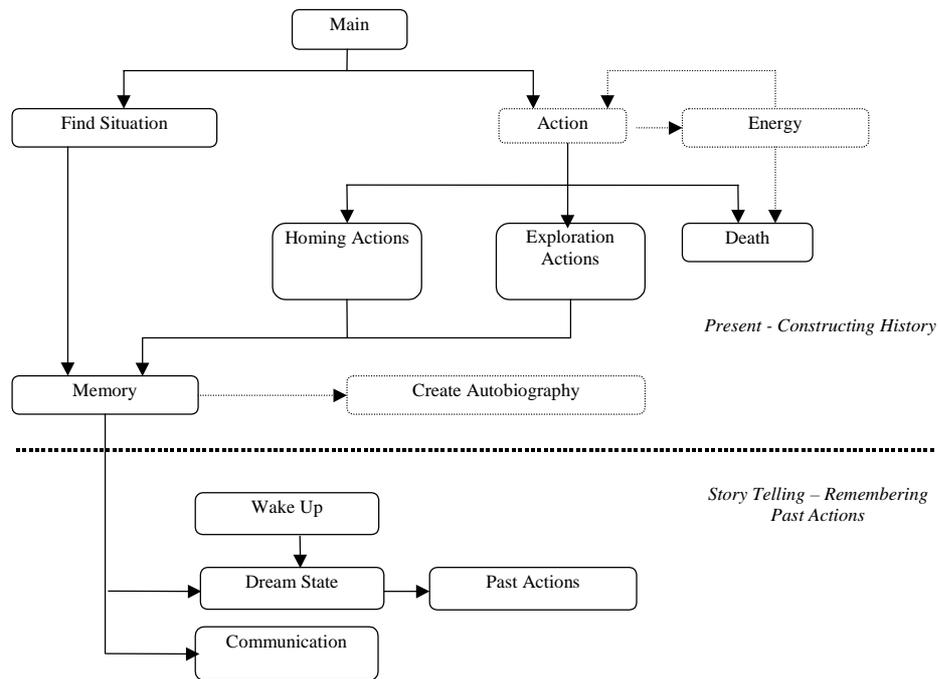


Figure 3: Graphical representation of the overall control program structure. The reactive (top) and post-reactive (bottom) control parts are shown. See text for more explanation.

<S, A, S'>

<13, 1, 19> <19, 2, 12> <12, 2, 12> <12, 2, 12> <12, 2, 12> <12, 2, 12>
 <12, 2, 12> <12, 2, 13> <13, 1, 13> <13, 1, 13> <13, 1, 19> <19, 2, 13>
 <13, 1, 19> <19, 2, 13> <13, 1, 13> <13, 1, 19> <19, 2, 19> <19, 2, 13>

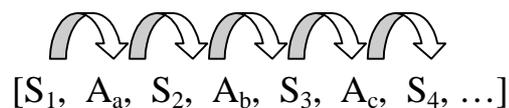


Figure 4: Representaton of Situation-Action-Situation triplets that form the core of the agent's story-representation and story recollection. A short part of an agent's autobiography are shown (read from left to right and top to bottom row). The integer values that characterise actions and situations represent classifications of sensor data from the infra-red and light sensors, cf. figure 1. A unique integer situation number is assigned to each situation combining both the light and proximity measures. Actions are defined in terms of wheel speed.

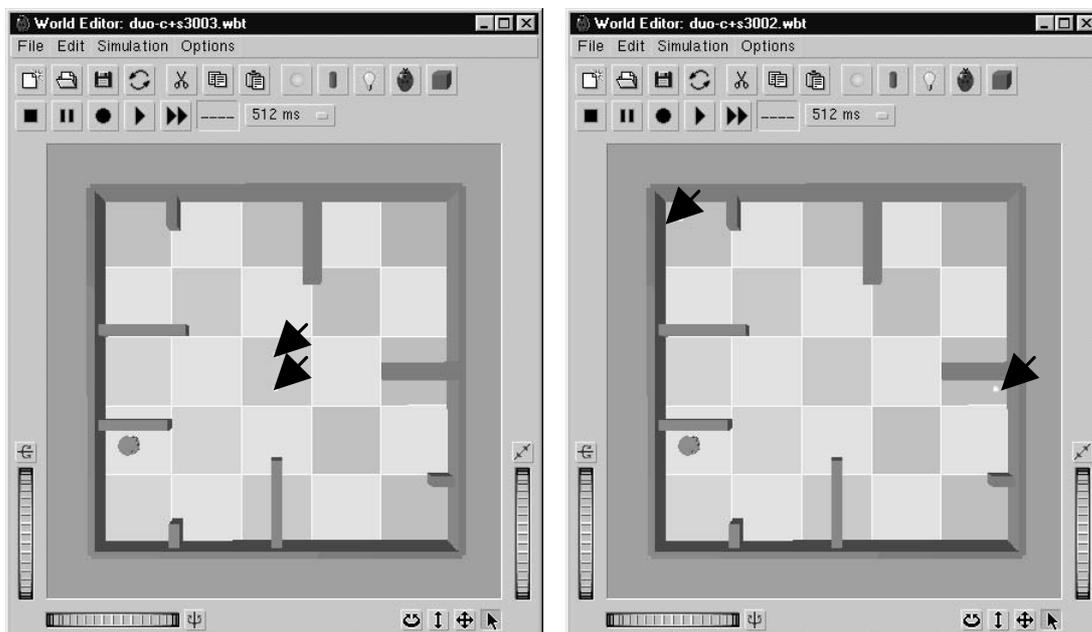


Figure 5: Environments used for test series A (right) and B (left). Arrows indicate positions of light sources. The colours of the "tiles" are a default configuration and serve no particular purpose in the experiments described. The ground is modelled as a flat surface, surrounded by stationary walls. The circular object in the left corner of the environments shows the robot at its starting location.

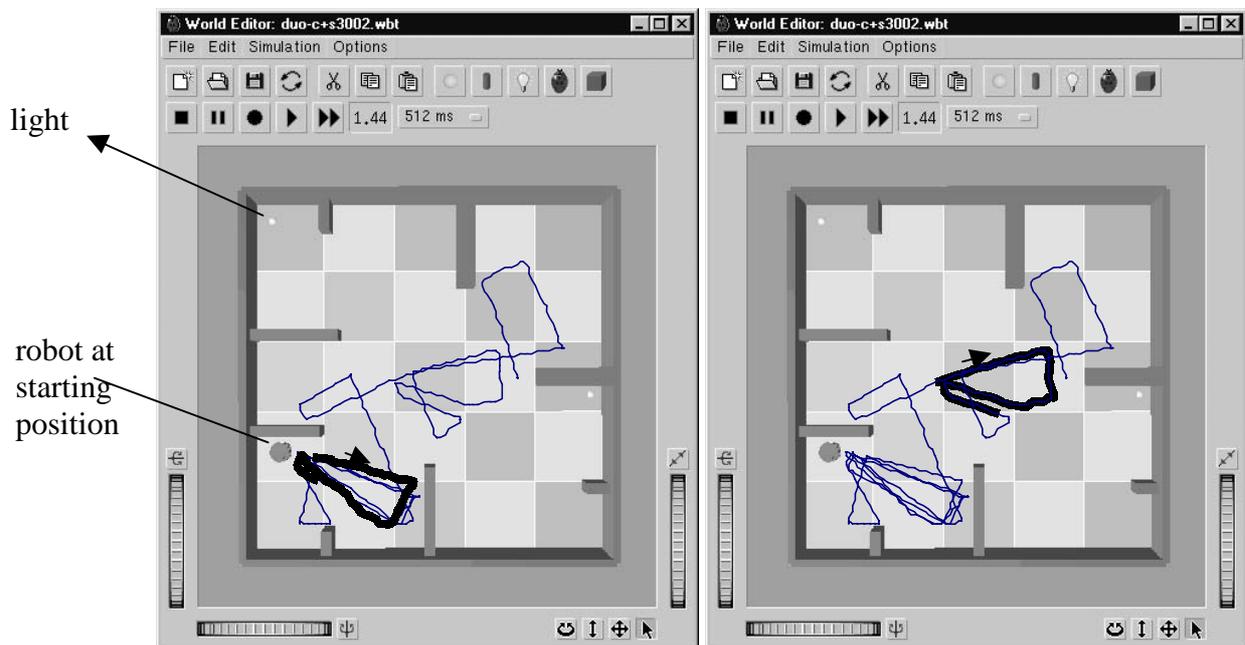


Figure 6: Comparison of trajectories produced by reactive (sensory-driven) versus post-reactive (memory-driven) control (cf. top and bottom part of figure 3), showing an example run of a robot remembering a sequence of actions. The left and right figure both refer to the same run (the same autobiography). The part of the trajectory that is a result of remembered actions is highlighted in the right figure. After examination of the robot's previous history (autobiography), the remembered trajectory part was identified, highlighted in the left figure. Arrows indicate the robot's direction of movement.

6 Results

Table 1 shows lifetimes for experiments A1, A2, B1 and B2. Figure 7 compares trajectory patterns for five runs from experiments A1, A2, B1 and B2.

Reactive	Post- Reactive	Post-reactive	Reactive
Environment A		Environment B	
8038	7459	3190	2200
4132	4109	6661	2200
4006	6845	3695	2200
4426	11470	3675	2200
6442	27048	3331	2200
4384	4203	3266	2200
10117	5755	3546	2200

Table 1 - Comparison of the lifetime of the agent with reactive versus post-reactive (memory-driven) controllers.

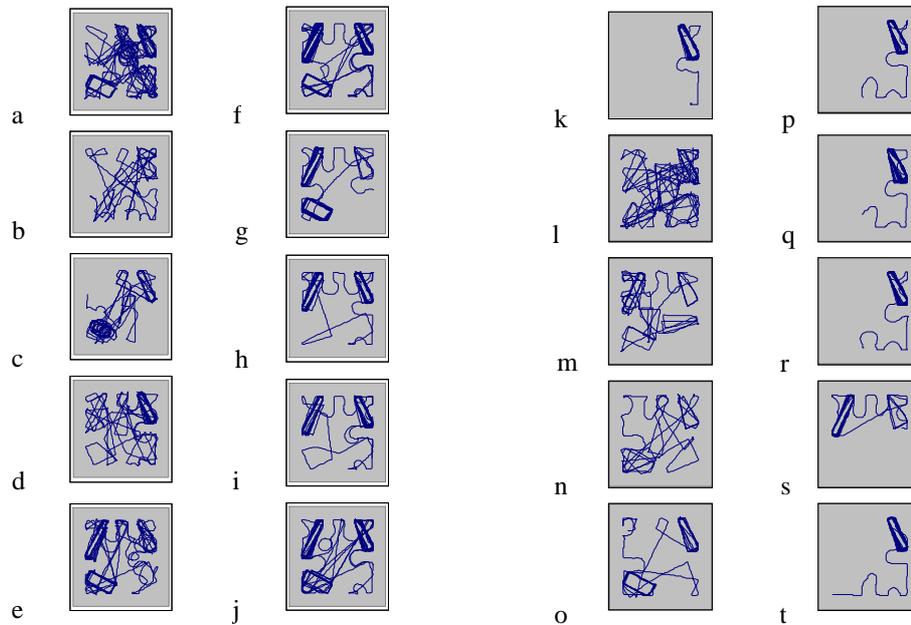


Figure 7: Environment A (a-j): comparison post-reactive (a-e) vs. reactive (f-j) control. Lifetimes a: 11470, b: 4109, c: 4203, d: 6845, e: 7459, f: 10117, g: 4132, h: 4384, i: 4426, j: 6442. Environment B (k-t): comparison post-reactive (k-o) vs. reactive (p-t) control. Lifetimes k: 3190, l: 6661, m: 36953, n: 3675, o: 3331, p-t:2200.

7 Discussion and Outlook

The comparison of the lifetimes shows no significant differences between experiments A1 and A2, i.e. the robot had about the same “life-span” in the purely reactive, as well as in the post-reactive mode. *Thus, in this particular environment the reactive controller performs as well as the post-reactive controller.* This is not surprising since, as described above, the robot's reactive controller was well adapted to survival in environment A.

However, if we analyse the trajectory patterns, we identify significant differences. In the purely reactive mode the robot either explores or homes (follows the walls), in a very fixed, stereotypical pattern. In the post-reactive mode the robot covers much more of the environment, in particular the central area (indicated by the number of crossings through the centre point).

Experiments B1 and B2 in environment B show that the robot lives significantly longer with a post-reactive architecture. The reactive architecture was designed so that in the homing phase the robot was looking for lights (recharging places) along the walls. This behaviour was no longer adaptive in environment B where the lights were located at a different location, a location the robot would usually never encounter with the purely reactive architecture. The robot died after 2200 simulations steps in all seven runs which means that it never encountered the charging station. The post-reactive architecture led the robot to explore a wider area so that its chances to encounter the charging station increased.

The experiments confirm our working hypothesis, namely that story-telling can provide an advantage to an autonomous, autobiographic minimal agent, i.e. an agent not knowing "what" it is remembering, e.g. whether the remembered actions are good or bad in the current situation.

However, further experiments under different experimental conditions are necessary. For example, the experiments we described above were controlled by fixed (experimentally determined) parameters (e.g. the length of the memory array) that need to be investigated further. However, it seems that a minimalistic architecture for story-telling can give an autonomous agent an adaptive advantage not necessarily by the meaning of what is remembered, but by the simple effect that remembering (and expressing these memories) provides sufficient "noise", i.e. interrupts the agent's normal (reactive) "routine", in our case leads the agent to explore a larger area of its environment than it would normally encounter. We can speculate that minimal mechanisms in the "first story-telling animal" survived because the animal was better adapted to a dynamic environment, story-telling increased its survival chances. In further experiments with minimal agents, and groups of such agents exchanging stories, under different experimental constraints, the investigation of what "meaning" can mean to an autonomous robot, and how we can progress from a day-dreaming to a genuinely "story-telling" agent will be one of the most challenging future studies along the framework presented in this paper. As mentioned above, the computational framework suggested in this paper studies *minimal* agents, results can therefore not directly be applied to humans who are sophisticated (Type IV) story-tellers.

At present we do not know if other animal species tell stories and what kind of stories they tell. Dolphins which possess a sophisticated communication system (e.g. Tyack 2000) are good candidates. However, we cannot expect that we as human beings will be able to understand the stories that non-human animals tell, given their different kind of embodiment and ecological niche. Thus, the line of investigation that this paper points out might one day result in robotic story-tellers that communicate with each other in terms of stories in "meaningful" ways (from the robots' points of view), even if humans (from an observer point of view) will not be able to understand the stories. However, if robotic agents acquire their story-telling skills in interaction with humans, and not with other robots (e.g. robots as described in (e.g. Brooks et al. 1999)), we can speculate that the stories told by robots then might be similar to the stories we tell, not from a phenomenological point of view, but from the point of view of shared experiences by being a member of a (cross-species) society.

Results of research into narrative for autonomous agents, along the lines motivated in this paper, can potentially contribute to research into *socially intelligent robots* that can interact with humans via narratives. We expect that contributions are likely to be on a conceptual level, not on the level on actual robotics controllers that require far more sophistication than the simple controllers investigated in this study. Because of the importance of stories in understanding and communication for human beings, robots equipped with such skills can provide a considerably more "natural" (cognitively adequate) interface to humans than standard interfaces. Speech interfaces have been developed to the degree that they can be realistically used in human-computer and also human-robot communication. Simple commands, or even phrases, triggered by a successful match of input data or internal processes, can be employed. However, as with any machines communicating with humans in natural language, a major bottleneck today is that systems usually have no genuine *understanding* of the words or sentences they are producing or parsing. Our bottom-up approach towards narrative autonomous agents is not presupposing any external definition of *meaning* or *interpretation* of stories, it stresses the interpretation of "meaning" as *meaning for a particular agent*, evaluated from its own (historical) perspective (Nehaniv 1999). "Understanding" in this framework then means that the agent's stories are grounded in its own experiences rather than imposed by a human designer. Human-robot communication in this sense is about negotiating meaning and developing means of communication based on socially constructed and shared meaning.

Acknowledgements. Thanks to Paul Brna and two anonymous referees whose comments helped improving previous versions of this paper. We like to thank Olivier Michel and Stéphane Magnenat for their kind permission to use and adapt some of their program routines in the development of the robot controllers. The project *Memory-Based Interaction in Autonomous Social Robots* is supported by a grant of The Nuffield Foundation NUF-NAL. Joint work with Chrystopher L. Nehaniv helped developing the theoretical basis for this research in previous joint publications. Steven Coles developed the initial software for story-telling robots. Kerstin Dautenhahn extended the software in order to run the experiments, and developed the ideas presented in this paper.

References

- Arkin, R. C. (1998), *Behavior-Based Robotics*, MIT Press.
- Aylett, R. (1999), Narrative in Virtual Environments – Towards Emergent Narrative, *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01, pp. 83-86.
- Billard, A. & G. Hayes (1999) DRAMA, a connectionist architecture for control and learning in autonomous robots, *Adaptive Behavior* 7(1), pp.35-64.
- Braitenberg, V. (1984), *Vehicles: Experiments in Synthetic Psychology*, MIT Press, Cambridge, Massachusetts.
- Brooks, R.A. (1986), A robust layered control system for a mobile robot, *IEEE Journal of Robotics and Automation* 2(1), pp. 14-23.
- Brooks, R. A. (1999), *Cambrian Intelligence : The Early History of the New Ai*, MIT Press.
- Brooks, R. A., C. Breazeal, M. Marjanovic, B. Scassellati, & M. M. Williamson (1999), The Cog Project: building a Humanoid Robot, In: *Computation for Metaphors, Analogy and Agents*, Springer Lecture Notes in Artificial Intelligence, Volume 1562, Ed. C.L. Nehaniv, pp. 52-87.
- Brothers, L. (1990), The social brain: A project for integrating primate behavior and neurophysiology in a new domain, *Concepts in Neurosciences* 1(1), pp. 27-51.
- Brothers, L. (1997), *Friday's Footprint: How Society Shapes the Human Mind*, Oxford University Press.
- Bruner, J. (1987), *Actual Minds, Possible Worlds*, Cambridge: Harvard University Press.
- Bruner, J. (1990), *Acts of Meaning*, Cambridge: Harvard University Press.
- Bruner, J. (1991), The Narrative Construction of Reality, *Critical Inquiry* 18(1), pp. 1-21.
- Coles S. J. & K. Dautenhahn (2000), A Robotic Story-Teller, *Proc. SIRS2000*, 8th Symposium on Intelligent Robotic Systems, The University of Reading, England, 18-20 July 2000.
- Conway, M. A. (1996), Autobiographical knowledge and autobiographical memories, In D. C. Rubin (Ed.), *Remembering Our Past. Studies in Autobiographical Memory*, Cambridge University Press, pp. 67-93.
- Dautenhahn, K. (1995), Getting to know each other - artificial social intelligence for autonomous robots, *Robotics and Autonomous Systems* 16, pp 333-356.
- Dautenhahn, K. (1996), Embodiment in Animals and Artifacts, *Proc. AAAI FS Embodied Cognition and Action*, AAAI Press, Technical report FS-96-02, pp. 27-32.
- Dautenhahn, K. (1998), Story-Telling in Virtual Environments, Working Notes *Intelligent Virtual Environments, Workshop at the 13th biennial European Conference on Artificial Intelligence (ECAL-98)*.
- Dautenhahn, K. (1999), Embodiment and Interaction in Socially Intelligent Life-Like Agents, In: C. L. Nehaniv (ed): *Computation for Metaphors, Analogy and Agent*, Springer Lecture Notes in Artificial Intelligence, Volume 1562, Springer, pp. 102-142.
- Dautenhahn, K. (1999b), 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.
- Dautenhahn, K., Stories of Lemurs and Robots - The Social Origin of Story-Telling, In M. Mateas & P. Sengers (Eds.), *Narrative Intelligence*, John Benjamins Publishing Company, forthcoming.
- Dautenhahn, K. & Aude Billard (1999), Studying robot social cognition within a developmental psychology Framework, *Proc. Eurobot99, Third European Workshop on Advanced Mobile Robots*, September 1999, Switzerland, pp. 187-194.
- Dautenhahn, K. & C. Nehaniv (1998), Artificial Life and Natural Stories, *Proc. of the Third International Symposium on Artificial Life and Robotics (AROB III)*, pp. 435-439, 1998.
- Dennett, D. C. (1989), The Origins of Selves, *Cogito* 3, pp. 163-73, Autumn 1989. Reprinted in Daniel Kolak and R. Martin, Eds., *Self & Identity: Contemporary Philosophical Issues*, Macmillan, 1991.
- Epstein, J. M. & R Axtell (1996), *Growing Artificial Societies: Social Sciences from the Bottom Up*, MIT Press.
- Franz, M. O. & H. A. Mallot (2000), Biomimetic Robot Navigation, *Robotics and Autonomous Systems* 30, pp. 133-153.
- Glenberg, A. M. (1997), What memory is for, *Behavioral and Brain Sciences* 20(1), pp. 1-55.
- Grand, S. & D. Cliff & A. Malhotra (1997), Creatures: Artificial Life Autonomous Software Agents for Home Entertainment, *Proc. First International Conference on Autonomous Agents (Agents '97)*, ACM Press, pp. 22-29.
- Kuipers, B. J. & T. S. Levitt (1988), Navigation and Mapping in Large-Scale Space, *AI Magazine*, pp. 25-43, Summer 1988.

To appear in the Journal of Artificial Societies and Social Simulation, JASSS,
<http://jasss.soc.surrey.ac.uk/JASSS.html>

- Kuipers, B. J. (1982), The "Map in the Head" Metaphor, *Environment and Behavior* 14(2), pp. 202-220.
- Kuipers, B. J. (1987), A qualitative approach to robot exploration and map learning, In *AAAI Workshop Spatial Reasoning and Multi-Sensor Fusion (Chicago)*, pp. 774-779.
- Langton, C. (Ed.) (1995), *Artificial Life: An Overview*, MIT Press, 1995.
- Machado, I., C. Martinho, & A. Paiva (1999), Once upon a Time, *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01, pp. 115-119.
- Mataric, M. (1992), Integration of Representation Into Goal-Driven Behavior-Based Robots, *IEEE Transactions on Robotics and Automation*, 8(3), Jun 1992, pp. 304-312.
- Mateas, M. & P. Sengers (1999), Narrative Intelligence, *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01, pp. 1-10.
- Mateas, M. & P. Sengers (Eds.), *Narrative Intelligence*, John Benjamins Publishing Company, forthcoming.
- Michaud, F. & M. J. Mataric (1999) Representation of behavioral history for learning in nonstationary conditions, *Robotics and Autonomous Systems*, 29, pp. 187-200.
- Michaud, F. & M. J. Mataric (1998), Learning from History for Behaviour-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, pp. 335-354.
- Nehaniv C. L. & K. Dautenhahn (1997), Semigroup Expansions for Autobiographic Agents, *Proc of the First Annual Symposium on Algebra, Languages and Computation*.
- Nehaniv, C. L. & K. Dautenhahn (1998), Embodiment and Memories - Algebras of Time and History for Autobiographic Agents, *Embodied Cognition and AI Symposium. Published in Proc. of 14th European Meeting on Cybernetics and Systems Research*, Editor: Robert Trappl, pp. 651-656.
- Nehaniv, C. L. (1997), What's Your Story? - Irreversibility, Algebra, Autobiographic Agents', *Proc. Socially Intelligent Agents*, AAAI Fall Symposium 1997, AAAI Press, Technical Report FS-97-02, , pp. 150-153.
- Nehaniv, C. L. (1999), Narrative for Artifacts: Transcending Context and Self, *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01, pp. 101-104.
- Nehaniv, C. L., K. Dautenhahn, & M.J. Loomes (1999), Constructive Biology and Approaches to Temporal Grounding in Post-Reactive Robotics, In G. T. McKee & P. Schenker, Eds., *Sensor Fusion and Decentralized Control in Robotics Systems II* (September 19-20, 1999, Boston, Massachusetts), Proceedings of SPIE Vol. 3839, pp. 156-167.
- Nelson, K. (1993), The psychological and social origins of autobiographical memory, *Psychological Science* 4(1), pp. 7-14.
- Nelson, K., Ed. (1986), *Event knowledge: structure and function in development*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pfeifer, R. & C. Scheier (1999), *Understanding Intelligence*, MIT Press.
- Quick, T., K. Dautenhahn, C. Nehaniv & G. Roberts (1999), 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, Liège, Belgium.
- Schank, R. C. & R. P. Abelson (1977), *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*, Erlbaum, Hillsdale, NJ.
- Sengers, P. (2000), Narrative Intelligence, *Human Cognition and Social Agent Technology*, Ed. Kerstin Dautenhahn, John Benjamins Publishing Company, pp.1-26.
- Sharples, M. (1997), Storytelling by Computer, *Digital Creativity* 8(1), pp. 20-29.
- Tani, J. (1996), Model-based Learning for Mobile Robot Navigation from the Dynamical Systems Perspective, *IEEE Trans. on System, Man and Cybernetics Part B (Special Issue on Robot Learning)*, Vol.26 (3), pp. 421-436.
- Tulving, E. (1983), *Elements of episodic memory*, Oxford: The Clarendon Press.
- Tulving, E. (1993), What is episodic memory?, *Current Directions in Psychological Science* 3, pp. 67-70.
- Turner, M. (1996), *The Literary Mind*, Oxford University Press.
- Tyack, P. L. (2000), Dolphins whistle a signature tune, *Science* 289, pp. 1310-1311.
- Umaschi Bers, M. & J. Cassell (2000), Children as Designers of Interactive Storytellers "Let me tell you a story about myself", in *Human Cognition and Social Agent Technology*, chapter 3, John Benjamins Publishing Company, edited by K. Dautenhahn, pp. 61-83
- Vygotsky, L. (1986), *Thought and Language*, Cambridge: The MIT Press.
- Weizenbaum, J. (1965), ELIZA- a computer program for the study of natural language communications between man and machine, *CACM* 9(1), pp. 36-45.
- Wertsch, J. W. (1985), *Vygotsky and the social formation of the mind*, Harvard University Press.

To appear in the Journal of Artificial Societies and Social Simulation, JASSS,
<http://jasss.soc.surrey.ac.uk/JASSS.html>

Wyer, R. S., Ed. (1995), *Knowledge and Memory: the Real Story*, Lawrence Erlbaum Associates,
Hillsdale, New Jersey