Narrative for Artifacts: Transcending Context and Self

Chrystopher L. Nehaniv
Interactive Systems Engineering
University of Hertfordshire
College Lane
Hatfield Herts AL10 9AB
United Kingdom

E-mail: C.L.Nehaniv@herts.ac.uk
WWW: http://www.cs.herts.ac.uk/∼nehaniv/welcome.html

Abstract

We discuss the importance of narrative intelligence (story-awareness, story-telling, historical grounding) in regard to an agent's transcendence of its immediate local temporal context to create a broad temporal horizon in which the experience and future of the agent can be accounted for, together with the advantage that narrative provides to sociality by making the experience of others available without the risk of having to undergo the experience for one's self. Concepts and consequences for the design of artifacts are surveyed, together with a brief description of a formal algebraic framework affording support for narrative grounding.

What's a Story For?

We address what it is about narrative that makes it worthwhile for natural and artificial agents. Interest in narrative in literature, cultural studies, psychology and the arts is of course much older than in Artificial Intelligence (AI) - in some cases ancient - and has necessarily focused primarily on human notions of experience and communication, and on human methods of structuring narrative flow in discourse, or other media such as film and books. This primarily natural human-based, or anthropocentric, viewpoint on narrative has now been extended metaphorically in various ways to agents.

The current widespread interest in narrativity and narrative intelligence in artificial intelligence (e.g. (Schank 1990)), in the design of artificial agents (e.g. (Sengers 1998; Elliott *et al.* 1998)), and as an important dimension in the consideration of system design (e.g. (Don 1990)) is a reflection of the importance of temporal structure and history in the experience of human agents in terms of dealing with the world. Such temporal structure has been or can be applied in allowing agents to take advantage of mechanisms used in natural historically grounded systems, and for making systems friendly to beings such as humans who possess narrative intelligence or expect it in artificial agents. Narrative aspects have been key in applications for both interfaces (e.g. (Don 1990)) and entertainment (e.g. (Lester & Stone 1997)), among many other areas, but have even been used in such applications as support for system development in software engineering, e.g. in scenario generation for requirements validation (Minocha *et al.* 1997; Maiden *et al.* 1998).

The topic of this paper, and related work of the author (among others), is to identify and formalize a notion of narrative for biological systems, agents and artifacts that may serve to clear the air to some degree on what is key in narrativity. By identifying what appear to be crucial features of narrative phenomena, this is intended to provide a useful framework for researchers and artifact designers in which to consider narrative. Some key concepts identified are *rigorous notions of meaning and information for agents and observers*. In contrast to most formal approaches, the approach here is centered on the structural coupling of agents with their environments, and the interaction dynamics between agents (whether biological, hardware, or software). Analyses along the lines of theoretical framework described here may be useful in the construction of narrative agents and artifacts, in the study of human and other animal narrative systems, as well as in related areas such as the study of imitation (Nehaniv & Dautenhahn 1998b) and emotion (Elliott *et al.* 1998; Rolls 1999; Nehaniv 1998; 1999 submitted).

Extrasensory Channels

A feature of narrative, but not necessarily of all forms of communication, is that *narrative provides an 'extrasensory' channel by which an agent may obtain meaningful information* (see below) *to modulate or guide its immediate or future behavior*. Related means in which this may occur include *memory* and *remembering* (often also involving narrative structure), and, with generally smaller temporal scope, *emotions*.

Temporal Horizons: Scope and Transcendence

Heidegger (Heidegger 1972) saw the state of man as being as situated in the Now, being there in the imminence of the Future in relation to the impinging Past. This 'temporal horizon' is extremely broad in humans,

apparently greatly so compared to other animals. The great temporal scope is evidenced by such of our emotions as hope and regret, our concern with planning for future actions and story-telling about past or imagined events (Nehaniv 1998; 1999 submitted).

Wittgenstein (Wittgenstein 1968) gave the famous, and supposedly absurd, example of a dog who expects his master to return the day after tomorrow. The perceived absurdity of this example is based on our (probably correct) attribution of much smaller scope to a dog's ability to envision temporal structure: although we can accept that that the dog might indeed expect his master's return "any time now", "soon", and perhaps even "tomorrow when the sun comes up again", the longer period in Wittgenstein's example implied to be within the scope of the dog's capacity for comprehension seems excessive, although we have no proof of this.

The vast temporal horizon of humans means that humans will tend to deal with interaction in a way that makes narrative sense, and may anthropomorphically expect their technological agents to do so. Affect and narrativity thus intertwine with each other and are relevant to issues of human-tool interaction and relations over time as well as to human-human interaction whether face-to-face or via networked media. But such 'extrasensory' data from narrative may help artificial agents to escape from the present and a rigid perception-action cycle as well.

Three Areas for Historical Grounding

The areas of temporal grounding for artifacts, in the interaction of humans and their tools can be divided into three areas:

(a) recognition of narrative structure: recognizing that others (humans, animals, agents) have a past history and using this in manner that allows present behaviour to be understood and future behaviour to be predicted to some extent. Relevant terms are *story-awareness* and *biographical reconstruction*. Related terms are: *projection of experience, 'hearing a story', detecting intention movements, attributing a history, learning from experience of others, literature*.

(b) expressing narrative structure: the communication of meaningful information regarding past or possible future events, i.e. events removed from the Now. Relevant terms are *story-telling, narration, explaining context, describing motivations and plans, teaching, communication*.

(c) having a narrative structure: the dependence and reliance of the agent on temporal dynamics in interacting with its world Relevant terms are *remembering, scheduling, planning, autobiography, autobiographic agent, historical grounding*. Peripheral terms are *intention, drives, emotions, moods, recording, state, development, trace, learning, growth*.

The terms *autobiographic agent* and *biographic reconstruction* were introduced by (Dautenhahn 1996), with formal development in (Nehaniv 1997; Nehaniv & Dautenhahn 1998c; 1998a).

Narrative and Design

Current software and agent technology has barely touched these three areas, but we believe they will be essential in freeing agents from social as well as temporal isolation. Moreover, in interaction with humans or in mediating human-human interaction over networked media, it is desirable to take into account that humans are temporally grounded, narratively intelligent beings. Their evolutionary heritage leads them to expect that the actions of others are embedded in a context of past history and future events. Software and robotic agents that do not respect the narrative grounding of humans may seem bizarre to them, may disappoint them, or may lead to cognitive calluses in human users that result in psychological and interpersonal changes in the course of long term interaction with technology as adaptations to it. Without support for narrative intelligence, such technology often fails to support human wholeness and does not optimize the human-tool relationship (Nehaniv 1999a). Together with affective issues, narrative and story-telling offer an important rich, undiscovered country loaded with opportunities and dangers for the designers of technology.

Streams of Meaning

The elementary nature of considerations here and below is intended to highlight important features of narrative, from a philosophical and cybernetic viewpoint, as a basis for rigorous study. In this exposition, we avoid a mathematical formulation, and proceed in a largely 'pre-formal' manner, that should also serve as preparation and motivation for rigorous case studies applying the mathematical framework of (Nehaniv 1997; Nehaniv & Dautenhahn 1998c; Nehaniv 1999b) to the illumination of narrative and interaction for artifacts.

We introduce below a notion of *meaning* for agents, which will be a pre-requisite for the notion of [meaningful] narrative.

Reactivity: Sensors and Actuators

Consider an agent from a first-person viewpoint, the experience of stimuli, especially through sensors, senses, and interaction with the environment via actuation, action and interaction in its situated (or embodied[1]). Events (including passage of time) trigger changes of its internal state.

Affect: Loosening the Grip of the Now

If the triggering stimuli are primary or secondary reinforcers in the sense of animal learning (i.e. stimuli which the agent will work either to obtain or avoid, either by its nature (primary) or after learning (secondary)), then the change of state is called by neurophysiologists an 'emotion' (e.g. (Rolls 1999)), and this state may modulate further behaviour. Similar to emotions, but of

[1]See (Quick *et al.* 1999) for the sense of 'embodied' and its degrees in both physical and non-physical settings.

longer term are 'moods', whereas 'drives' are analogs of emotion that result from internal state, perhaps in interaction with environment stimuli, to bring homeostatic parameters into preferred ranges. 'Taxis' and 'tropism' are names for reactive versions of such state changes connected directly with inflexible behaviours. The level of taxis or tropism is close to much of the basic level in behavioral robotics as developed in the school of Rodney Brooks.

The cost and reward of experience stimuli provide a single dimension (a kind of 'common currency' (Rolls 1999)) along which to evaluate the result or desirability of action, and the relative values of these costs and rewards (or 'pains' and 'pleasures') may be modulated by the current state of the agent. Most attempts to introduce 'emotion-parameters' into AI systems, can be seen as an attempt to solve the well-known *contextualization problem* of AI and robotics, for instance, to transcend simple reactivity by allowing the settings of the parameters to modulate behaviour, so as to respond appropriately to the given context or to provide a basis for learning (See (Nehaniv 1999 submitted)).

Meaningful Information

Information in the channels of sensing and actuating for an agent, including interaction channels that affect its state, as well a channel related to the general passage of time (which may be different in biological and artificial agents, indeed possibly cyclical, discrete, or continuous, and need not be correlated with time for other agents) are referred to as *meaningful* (Nehaniv 1999c; 1999b) for the agent. A channel may also be meaningful for an observer who attributes such changes to the agent, whether consciously or not (Nehaniv 1999c; 1999b).

Formal Narratives for the Self and Others

Sequences events may also be recorded in ad hoc, haphazard and opportunistic ways in real biological systems. This may range from state change in response to a history of stimulation, to neural changes, learning, development, and growth. They may be of small or large temporal scope.

Records of sequences of events and/or of their resulting state changes, encoded in systematic ways, are called *formal stories* or *narratives*. For a given agent or observer, they comprise an algebraic structure called a '*historical expansion*' in the the algebraic theory of abstract automata and semigroups (Nehaniv 1997; Nehaniv & Dautenhahn 1998c; 1998a). An element of such an expansion over a meaningful channel (in the sense of the preceding section) is then a *meaningful narrative*. Note that the qualifier 'meaningful' only makes sense in relation to the particular agent or observer involved in 'witnessing' the channel (see above). There may be several such agents and observers in any particular case.

A narrative in this formal sense simply consists of a systematic record of events and resulting state transitions for the particular agent. This record need not be an exhaustive specification, but may focus in a systematic [functorial] way on particular aspects of the events. In this way, the narrative is a historical record of experience either of an agent itself or attributed to one by an observer – of course, the narrative could possibly be an invented one rather than a result of actual experience.

Transcendence of the Present and Self

Such narratives may serve as a basis for memory, and in this case they may be considered stories about the self. Transmission, by whatever means, of such a formal narrative to another agent may be meaningful for the second agent. This implies that the state of the second agent can be affected by receiving the narrative. If the receiving agent has a similar enough body-environment coupling and the means to 'decode' or interpret the narrative, there is a possibility that the narrative could be useful to it. That is, the narrative could serve as something meaningful (in the sense above) to the recipient. Some ways in which this happens in real life, is that the agent is able to learn from the stories, or experience, of the first agent.

Both the cases of narrative for the self or for another provide possible streams of meaning for agents other than direct sensory ones, and moreover, the changes induced by 'hearing' a story need not be 'emotional' but may nevertheless serve to modulate behaviour. Yet they break the reliance of an animal or machine-learning agent on a need to experience the consequences (e.g. costs and rewards) of its own actions directly, either by making use of past experience of the self (remembering a story) or of another (narrativity). In this way, narrative can be viewed as at least as fundamental to intelligent behaviour, at least in humans (and possibly bonobos and dolphins?), as are emotion and affective response in the modulation of behavior in other animals and agents in escaping from mere reactivity.

From this perspective, narrative is another means not only to approach solution of the contextualization problem, but to transcend the first-person in learning by making use of the experience of another, a second-person (Nehaniv 1997; 1998; 1999c; 1999 submitted). The use of narrative in grounding intelligence is in this sense a second-person technique for building 'intelligent' artifacts grounded in a social world as well as by their structural coupling to the environment. Indeed, it has been argued that the nature of reality itself as experienced by humans and human cultures is an emergent effect of narrative interactions (Bruner 1991). (Such emergence of a shared reality is of course not arbitrary but must be grounded in the particular embodiments and interactions of the agents.) This may well also prove the case for important applications of artificial agents that interact with humans and with each other, and the formal methods overviewed here should provide at least one inroad.

References

Bruner, J. 1991. The narrative construction of reality. *Critical Inquiry* 18(1):1–21.

Dautenhahn, K. 1996. Embodiment in animals and artifacts. In *Embodied Action and Cognition: Papers from the AAAI Fall Symposium (Cambridge, Mass. - November 1996*, volume FS-96-02. AAAI Press.

Don, A. 1990. Narrative and the interface. In Laurel, B., ed., *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, MA. 383–391.

Elliott, C.; Brzezinski, J.; Sheth, S.; and Salvatoriello, R. 1998. Story-morphing in the affective reasoning paradigm: Generating stories semi-automatically for use with 'emotionally intelligent' multimedia agents. In Sycara, K. P., and Wooldridge, M., eds., *Proceedings of the Second International Conference on Autonomous Agents (New York, May 1998)*. ACM Press.

Heidegger, M. 1972. *On Time and Being*. Harper Torchlight Books.

Lester, J., and Stone, B. 1997. Increasing believability in animated pedagogical agents. In *Proceedings of the First International Conference on Autonomous Agents, February 1997*. ACM Press.

Maiden, N. A. M.; Minocha, S.; Manning, K.; and Ryan, M. 1998. CREWS-SAVRE: Systematic scenario generation and use. In *Third IEEE International Conference on Requirements Engineering. An IEEE Software Technology Transfer Conference, 6-10 April 1998), Colorado Springs, Colorado, USA*.

Minocha, S.; Maiden, N.; Ryan, M.; and Manning, K. 1997. A software tool for scenario generation and use. In *Third International Workshop on Requirements Engineering: Foundation for Software Quality, June 16-17, Barcelona, Catalonia*.

Nehaniv, C. L., and Dautenhahn, K. 1998a. Embodiment and memories - algebras of time and history for autobiographic agents. In Trappl, R., ed., *Cybernetics and Systems '98 (Proceedings of the 14th European Meeting on Cybernetics and Systems Research Symposium on Embodied Cognition and Artificial Intelligence (Vienna, Austria, 14-17 April 1998))*, volume 2, 651–656. Austrian Society for Cybernetic Studies.

Nehaniv, C. L., and Dautenhahn, K. 1998b. Mapping between dissimilar bodies: Affordances and the algebraic foundations of imitation. In Demiris, J., and Birk, A., eds., *Proc. European Workshop on Learning Robots 1998 (EWLR-7), (Edinburgh, Scotland - 20 July 1998)*.

Nehaniv, C. L., and Dautenhahn, K. 1998c. Semigroup expansions for autobiographic agents. In *Proceedings of the First Symposium on Algebra, Languages and Computation (30 October-1 November 1997, University of Aizu, Japan)*, 77–84. Osaka University.

Nehaniv, C. L. 1997. What's your story? - irreversibility, algebra, autobiographic agents. In Dautenhahn, K., ed., *Socially Intelligent Agents: Papers from the 1997 AAAI Fall Symposium (November 1997, MIT, Cambridge, Massachusetts)*, volume FS-97-02, 150–153. American Association for Artificial Intelligence Press.

Nehaniv, C. L. 1998. The first, second and third person emotions: Grounding adaptation in a biological and social world. In Cañamero, D.; Numaoka, C.; and Petta, P., eds., *From Animals To Animats: Fifth International Conference of the Society for Adaptive Behavior (SAB'98) Workshop on Grounding Emotions in Adaptive Systems (Zürich, Switzerland - 21 August 1998)*, 43–47.

Nehaniv, C. L. 1999a. Emotion and story-telling: Cognitive technology considerations in networking affectively and narratively grounded minds. In *Proceedings of the Third International Conference on Cognitive Technology (CT'99 - San Francisco/Silicon Valley, 11-14 August 1999)*. MIND Laboratory.

Nehaniv, C. L. 1999b. Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control/Intelligent Systems and Semiotics, ISIC/ISAS'99 - September 15-17, 1999 Cambridge, Massachusetts*.

Nehaniv, C. L. 1999c. The second person - meaning and metaphors. In *Computation for Metaphors, Analogy and Agents*, volume Lecture Notes in Artificial Intelligence Vol. 1562. Springer Verlag. 380–388.

Nehaniv, C. 1999 (submitted). Constructive biology of emotion systems: First and second person methods for grounding adaptation in a biological and social world. *Cybernetics and Systems* (Guest Editors: D. Cañamero, P. Petta, C. Numaoka).

Quick, T.; Dautenhahn, K.; Nehaniv, C. L.; and Roberts, G. 1999. On bots and bacteria: Ontology-independent embodiment. In *Proc. European Conference on Artificial Life (ECAL'99)*.

Rolls, E. T. 1999. *The Brain and Emotion*. Oxford University Press.

Schank, R. 1990. *Tell Me a Story: A New Look at Real and Artificial Memory*. Scribner, New York.

Sengers, P. 1998. *Anti-Boxology: Agent Design in Cultural Context*. Ph.D. Dissertation, Carnegie Mellon University.

Wittgenstein, L. 1968. *Philosophical Investigations, (Philosophische Untersuchungen), German with English* translation by G. E. M. Anscombe, 1964. Basil Blackwell, Oxford, reprinted 3rd edition.