What's Your Story? — Irreversibility, Algebra, Autobiographic Agents

Chrystopher L. Nehaniv
Knowledge Engineering, Cybernetics and Software Systems Group
Graduate School of Computer Science and Engineering
University of Aizu
Aizu-Wakamatsu City, Fukushima 965-80 Japan
nehaniv@u-aizu.ac.jp www.u-aizu.ac.jp/~nehaniv

Abstract

Biological systems exist under circumstances of irreversibility that make them fundamentally different from inanimate matter (distinguishing them for instance as subject to development, damage and death). Individual historical memory and story-telling are capacities useful in coping with this irreversibility. Such mechanisms may for embodied agents contribute to — or may be even be necessary for — competence in social intelligence.

Structuring historical memories connects to an area of algebra called global semigroup theory, which allows one to construct expressions in 'algebras of time' that can support recording events of fundamental significance to an agent. This may include records of internal changes (e.g. in motivational, goal, emotional, cognitive, body image, perceptual states, knowledge of other entities) as well as external changes (location, objects or other agents in the environment, physical conditions of local environment) that provide grounding for the representation. The models of change and time selected can be different depending on what is important to the agent.

Communicating knowledge of other minds could be modelled in this framework by the passing of algebraic expressions (in a historical 'expanded' semigroup) which record changes in the life-long learning or experience of another agent. This would correspond to revealing one's autobiography or, more generally, to telling a story.

Introduction and Overview

This paper is concerned with the representation of histories as autobiographies of social agents using ideas from algebra. It suggests an approach for such representations which would be amenable to algebraic and algorithmic manipulation, and provide methods for constructing a social agent's autobiography while selecting only certain information as relevant to the agent. This perspective provides a framework in which one can approach the problems for embodied agents of how (1) to reconstruct or use one's own autobiography, (2) to listen to a story in order to reconstruct a portion of another agent's biography, and (3) to represent the history of or the 'mind' of another agent.

First I present the notion of mediating the interface between humans, other living organisms, or embodied agents and the world via material and cognitive tools. Algebraic methods can algorithmically generate a class of cognitive tools, useful in providing coordinate systems affording understanding. In particular, these methods are useful in capturing irreversible aspects of life that arise for living systems.

This leads to suggestions for methods of algebraically constructing/manipulating histories for autobiographical agents. Various choices are possible for what constitutes relevant information in a history, and these correspond to alternative algebraic expansions for events in time. Grounding a historical representation in embodied agents to their sensory capacities, physical bodies and the environment where they are situated[1] gives meaning to such algebraic expressions of history for social agents. Portions of history (stories) may be transmitted between social agents, comprising autobiographies from 'other minds', which when multiplied in an appropriate algebra yield a representation of the other's biography.

Our (extensible) bodies (and minds): Our cyborg selves

Extending the viewpoint of D. J. Haraway (Haraway 1991), I argue (Nehaniv 1997) that humans — indeed biological systems — have always extended themselves via 'tools' and hence are *cyborgs*. Moreover, we integrate into ourselves various abstract tools (such as coding, language and arithmetic formalisms) of a nonphysical nature. Prosthetics that we may choose to use dynamically to augment our bodies (e.g. a pair of eyeglasses, a hammer, a microscope) exist also at the more abstract level of cognitive tools (e.g. language, cultural knowledge, arithmetic, cartesian coordinates). These abstract tools are as 'artificial' as aircraft or wheelchairs and as much (or more so) a part of self as the fillings in our teeth.

[1]See the discussions of *situatedness* ("the world is its own best model") and *embodiment* ("the world grounds regress") in (Brooks 1995).

Mathematics shows that for any given circumscribed finite-state system and its transformations (whether reversible or not), one may obtain an appropriate coordinate system representation (Krohn, Rhodes, & Tilson 1968). Such an algebraic coordinatization may serve as a new mental lever for understanding the given system (Nehaniv 1996). Descriptions of the necessary techniques and examples including how the decimal expansion, encodings in clock faces, coordinate systems for understanding symmetries of regular geometric figures, and conservation laws in physics can be derived in this way may be found in (Nehaniv 1997).

Biological Systems and Irreversibility

Erwin Schrödinger (Schrödinger 1944) in his inquiry into the nature of living systems speculated that they are analogous to 'aperiodic crystals'. Such aperiodic crystalline properties have since been found in the double-helical molecule of heredity, deoxyribonucleic acid (DNA), which is at the same time both stable and regular while providing the basis for encoding genetic variability. Thermodynamically, living systems maintain themselves while causing a stream of order to come raining down on them, via their metabolisms, at the expense of an irreversible global increase in entropy in their surrounding environments.

Finally each living thing is doomed to lose its individual battle against entropy. Many irreversible events occur in the development of most biological systems.[2]

Development is largely irreversible, certainly very much so for humans and other mammals. Damage in the course of life may be mildly or extremely irreversible (scarring or the loss of a limb), as are other changes (getting one's adult teeth, development of secondary sex characteristics, learning to read, or metamorphosis from chrysalis to butterfly). And ultimately, death, as the loss of homeostasis and maintenance of the body's barriers against the increasing entropy of the environment, reveals a fundamental distinction between ourselves as embodied living beings and inanimate systems such as crystals, chemical compounds or planetary systems following computable trajectories.[3]

[2]It may be surprising the some events that would be irreversible for humans would not be so for certain other creatures. For example, in Porifera (sponges), cells of the organism passed through a sieve can regenerate new sponges (Gilbert 1994, pp. 28–29).

[3]The latter are amenable to formal description by differential equations in which one may recover previous states of the system by putting in, say, a negative number for a time parameter t. Such an approach is of course inadequate — by the uniqueness theorems for differential equation models — if used for instance to attempt to recover prior courtship behavior of a bird sitting outside your windowsill, or more generally in recovering previous states *when the current state could have resulted from any of several possible histories.*

Algebras of Time and History

Probably as early as Aristotle, philosophers realized: Time has the property that event c following the two successive events a then b is the same as b followed by c and prior to both a. Thus events in time satisfy the *associative law* (Nehaniv 1993): $(ab)c = a(bc)$. Structures satisfying this law are called *semigroups.* Unlike the time of differential equations (which fails to be expressive enough for living systems), associativity gives us a framework for myriad models of time, even if irreversible.

You can't go back.

For example, in multiplying in the semigroup of numbers: $3 \times 6 \times 0 \times 8 = 0 \times 15$. The result is zero, no matter what other factors were involved in the product. Knowing only the result (zero) does not allow us to recover any other factors. Similarly, considering two dead insects zapped by an electrical bug killer or two inhabitants of ancient Athens, one cannot recover their different histories or much else to distinguish them from their remains. The situation would be much improved if one of them kept a still extant diary (or interacted with someone who did).

We'll always have Paris: Capturing Memories

What exactly were you doing a year ago today? What were you wearing? What did you eat? Without films, photographs, historical records, knowledge of this is likely to be gone forever. However, some physically invisible facts from the past are recoverable since they are remembered. By keeping track of histories of a calculation or a life, much more than current state is available. For example, a history of the left hand side above is: 3, 18, 0, 0 while on the right we have 0, 0 as two stories of partial products when calculating from left to right. For a living organism, a record of special irreversible state transitions beginning with the start of life and ending with death provides a selective narrative: hatch from the egg, pupate, narrowly escape being eaten by a bird. In fact for any semigroup, the set of possible (in a fixed representation) histories themselves have an associative multiplication structure. Thus the set of possible histories itself comprises an expanded algebraic structure.[4]

Note that the decimal expansion of the real numbers also encodes a history of how to obtain the number. For example, 1.9999.. and 2.000.. describe very different paths to the number 'two': "Take one and nine-tenths and nine one-hundredths and ..." vs. "Start at two and take zero and zero and zero ..." It is worthwhile to remark that even here in arithmetic we have two histories for the same result.[5]

[4]The current state can be recovered from the history, which is thus *literally* an expansion of the present state.

[5]Also observe that the irrational numbers are exactly

Even in learning a simple vocabulary describing the actions in environment from another agent, a social robot can never have exactly the same observational perspective as its teacher, so it would be useful to remember at least enough history to attempt to transform the perspective of the teacher to that of the learner. A concrete example of the difficulties in an a-historical learning of such a vocabulary is presented in (Billard & Dautenhahn 1997).

Expansions: Different Views of Time.

The choice of what to remember in the history is by no means unique. For example in the products of numbers above, we get different histories if we record, the partial products from left-to-right, right-to-left (choosing a direction for time), or record all the factors (among many other possibilities).

Newtonian histories consist of records of all instances of physical time and associate events that occurred at each. Such a notion of historical time is somewhat bizarre from the perspective of living systems for whom only especially important events may be relevant, and so ultimately probably it will not be of great interest for them or for use in other embodied agents. This realization about individual relevance of events is closer to Freud's notion of intense or traumatic events as determinants of an individual's character (birth, typical event sequence yielding Oedipal complex, first sex, marriage, loss of parent, ...), which comprise the elements of a *Freudian history.*

Other notions of history may include records of reflections on the past events or future events (applicable in planning, but also in neurosis), and so on. These different sets histories all correspond to semigroup expansions used in algebra.[6]

The semigroup for possible events in the life-history of a robot or organism varies strongly with the structure of the entity and its sensory, perceptual and other capacities. Given a semigroup, expansions are uniform ways of enlarging the semigroup into another one whose elements include formulaic encodings of historical information about events possible in the semigroup. It is thus possible to expand various heterogeneous semigroup structures in the same way, e.g. Newtonian histories can be constructed for two physically radically divergent entities.

While global semigroup theory, and in particular expansions, are a recently developed area of mathematics that is only slowly finding its way into applications in the inanimate natural sciences, it seems clear that it provides a language in areas where traditional mathematics is inadequately mute. Namely, the sciences of the artificial, including design sciences, robotics, and biology (especially evolutionary and developmental biology and ecology), exhibit a crucial historical and contingent nature. Here the fact of designed, evolved or adapting interfaces to the world or to other entities or to portions of the self, which *could have been otherwise,* indicates the need for a mathematical language that can usefully formalize the notion of 'what happened'. We are not in Kansas anymore.

Autobiographic Agents.

I define the concept of an autobiographic agent as an embodied agent which dynamically reconstructs its individual 'history' (autobiography) during its lifetime.

-K. Dautenhahn (Dautenhahn 1997b)

Now instead of purely reactive systems controlling our synthetic social agents (whose 'mind' may be described as a single state with little analyzable structure and no historical information), the foregoing discussion reveals that algebra provides various options in replacing simplistic states by expanded ones which include history of the agent elaborated in a particular way. On-the-fly construction of the histories is possible for many of the expansions. A history can serve as an element of an algebraic structure, but also as an autobiographical portrait of what is important to the agent about itself and of how it got the way it is.

Depending on the available sensory/perceptual modalities, representation of motivation, goals, or 'emotional' state in a social robotic agent, one can choose an algebra for time and an expansion for history in which the agent could compute during its experience or re-construct from memories autobiographical information in an appropriate algebraic system. Furthermore, agents with varying capacities, endowments and experience may use different historical algebras.

Prospects: Stories of Other Minds

Such autobiographies could of course be communicated to a human or artifact social agent. Furthermore, the algebraic structure on histories means that one has algorithms to build-up computed histories from successive portions of history. That is, elements of a formal algebraic story may be multiplied by a receiving agent to yield a history of the sending agent. Such an approach may provide insight into the representation by social agents of other minds, which has been suggested as a fundamental pre-requisite for social intelligence; while autism is hypothesized by many researchers to indicate a failure to construct empathic representations of other minds.[7]

those whose historical (decimal) expansion representations never repeat in eternal cycles.

[6]The most well-known of these is the *Rhodes expansion,* which is useful in complexity theory and records irreversible transitions as one reads the events from most recent to longest ago, while essentially ignoring reversible events. See (Tilson 1976) for its construction, the multiplication of Rhodes histories, and mathematical applications.

[7]See (Dautenhahn 1997a) for a survey and thought-provoking discussion of the issues of other minds as related to socially intelligent agents, and some possible cog-

Other minds may or may not suffer from some pathology, e.g. *neurosis* can result from overactive attempts at constructing histories that take many past events and possible future events into the account; in the extreme case, the present could be lost from consideration in spite of the fact that the agent effects actions only in the present.

Implementing historical representation in autonomous agents (e.g. robots) and the challenge of usefully relating histories between agents of the same or varying kinds remains a future prospect. Forthcoming work of K. Dautenhahn and the author focuses on the realization of historical representations, motivation and emotional states, and their communication (story-telling) in appropriate coordinate systems between situated agents.

This paper reflects our early attempt at outlining how one might exploit algebraic methods in autobiographic agents. Technical details of the semigroups and appropriate historical systems based on them will depend on the characteristics of the agents who are to make use of them. In future work it should be possible, for example, (1) to test and compare the efficacy of different mechanisms for recording histories (different coordinate systems arising from various semigroup expansions of the agent's possible transitions) such as Newtonian, Freudian, Rhodesian, other or dynamically varying historical systems in embodied social agents, (2) to compare, for various problems, autobiographic agents to reactive a-historical agents who process only the immediate context and a (minimal) state to select their actions, (3) to allow evolutionary emergence of choices for how to select, encode and decode historically relevant information, (4) to study communication of histories as they are constructed (hearing an autobiography as it develops) between robotic agents which may be using the same or different representation for their own autobiographies as their social partners are using, (5) to build artificial societies of story-telling robots with similar problems to solve and to study the evolution of emerging cultural information and the agents' social construction of meaning, (6) to study possible social pathologies transmitted via story-telling in a culture of embodied socially intelligent agents, and (7) to study how story-telling influences polymorphism of social roles and ontogeny of communicating agents (see the discussion of (esp. third-order) *structural coupling* in (Maturana & Varela 1992)).

---

nitive mechanisms whose failure may be related to types of autism: Mechanism A provides for *empathic 'resonance'*, that is a degree of 'empathic, bodily re-experiencing' of the (*state* of) others. Mechanism B relies on this, but interprets not only 'the immediate situation of a human, but [also] his/her *biographical context*', which is usually not directly visible, but comprises important *re-constructed* aspects of the other individual's *history*. Note that from an algebraic viewpoint, mechanism B is analogous to an expansion (in the mathematical sense) of mechanism A.

Acknowledgements

I am indebted to Professor John L. Rhodes of the University of California at Berkeley for the deepest part of an unusual mathematical education. The realization that semigroup theory is related to history comes from him. It is pleasure also to acknowledge Dr. Kerstin Dautenhahn of the University of Reading for stimulating discussions and inspiration.

References

Billard, A., and Dautenhahn, K. 1997. Grounding communication in situated, social robots. In *Proc. "Towards Intelligent Mobile Robots" TIMR UK 97.* Technical Report Series of the Department of Computer Science, Manchester University.

Brooks, R. A. 1995. Intelligence without reason. In Steels, L., and Brooks, R., eds., *The Artificial Life Route to Artificial Intelligence*, 25–81. Lawrence Erlbaum Associates, Inc.

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

Dautenhahn, K. 1997b. The role of interactive conceptions of intelligence and life in cognitive technology. In Marsh, J. P.; Nehaniv, C. L.; and Gorayska, B., eds., *Proceedings of the Second International Conference on Cognitive Technology*, 33–43. IEEE Computer Society.

Gilbert, S. 1994. *Developmental Biology.* Sinauer Associates, Inc., 4th edition.

Haraway, D. J. 1991. A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, Cyborgs and Women: The Reinvention of Nature.* Routledge.

Krohn, K.; Rhodes, J.; and Tilson, B. 1968. Chapters 1,5,7-8. In Arbib, M., ed., *Algebraic Theory of Machines, Languages, and Semigroups.* Academic Press.

Maturana, H. R., and Varela, F. J. 1992. *The Tree of Knowledge: the Biological Roots of Human Understanding.* Shambala Publications, Inc., revised edition.

Nehaniv, C. L. 1993. The algebra of time. In *Proc. National Conf. Japan Soc. for Industrial and Applied Math.*, 127–128.

Nehaniv, C. L. 1996. Algebra & formal models of understanding. In Ito, M., ed., *Semigroups, Formal Languages and Computer Systems, RIMS Kyokyuroku*, volume 960, 145–154. Kyoto Research Institute for Mathematical Sciences.

Nehaniv, C. L. 1997. Formal models for understanding: Coordinate systems and cognitive empowerment. In *Proceedings of the Second International Conference on Cognitive Technology*, 147–162. IEEE Computer Society.

Schrödinger, E. 1944. *What is Life?* Cambridge University Press, reprinted 1992.

Tilson, B. 1976. Chapter XII: Complexity of semigroups and morphisms. In Eilenberg, S., ed., *Automata, Languages and Machines*, volume B. New York: Academic Press.