Introduction: Why "Epigenetic Robotics"?

Jordan Zlatev and Christian Balkenius

Background

During the last few years we have witnessed the mutual rapprochement of two traditionally very different fields of study: developmental psychology and robotics. This has come with the realization by large parts of the cognitive science community (e.g. Ziemke, in press) that true intelligence in natural and (possibly) artificial systems presupposes three crucial properties:

(a) the *embodiment* of the system;
(b) its *situatedness* in a physical and social environment;
(c) a prolonged *epigenetic developmental process* through which increasingly more complex cognitive structures emerge in the system as a result of interactions with the physical and social environment.

To designate this new field, we use the term *epigenesis* introduced in the field of psychology by Jean Piaget, the great 20th century developmentalist. The term was used to refer to such development, determined primarily by the *interaction* between the organism and the environment, rather than by genes. However, we believe that Piaget's emphasis on the importance of sensorimotor interaction needs to be complemented with what is just as (and perhaps more) important for development: *social* interaction, as emphasized by Lev Vygotsky, another important figure of 20th century psychology.

In the emerging field of *Epigenetic Robotics*, the interests of psychologists and roboticists meet. The former are in a position to provide the detailed empirical findings and theoretical generalizations that can guide the implementations of robotic systems capable of cognitive (including behavioral and social) development. Conversely, these implementations can help clarify, evaluate, and even develop psychological theories, which due to the complexity of the interactional processes involved have hitherto remained somewhat speculative.

With this in mind, we invited the submission of papers to the *First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, sponsored by the Communications Research Laboratory, Japan, and held in Lund, Sweden on September 17-18, 2001. Our vision was that this workshop would serve as a forum for sharing and discussing theoretical frameworks, methodologies, results and problems in this new interdisciplinary field.

We were most pleased to receive many original contributions. After reviewing, we could accept for presentation the 16 papers published in this volume together with those of three invited speakers: the animal psychologist Irene Pepperberg, the child psychologist Chris Sinha, and the 'robot psychologist' Tom Ziemke. In the following paragraphs, we briefly present these 19 papers, relating them to the theme of the workshop.

Learning and development

In an invited paper, **Irene Pepperberg** opens the discussion by highlighting an issue of crucial importance: What is the role of "desires and purposes", rich social interaction and even social awareness for acquiring complex communicative and cognitive competence? With a long experience of thorough and insightful experimentation with teaching the basics of language to Grey parrots, Pepperberg convincingly argues that these are indeed necessary prerequisites, while simple conditioning is inadequate for any kind of "advanced learning" to arise. The author furthermore suggests that the detailed study of learning and development in birds would help us both understand better these processes in human beings, and "improve the ability of non-living computational systems".

The capabilities of associative learning systems are also addressed by **Bridget Hallam**. She extends her earlier work on Halperin's Neuro-Connector model with simulations of behavioral chaining and the learning of new releasers. The simulations show that the model can learn new behaviors relatively quickly and may also become a starting point for more complex developmental processes.

Another type of associationistic learning is described by **Louis Hugues and Alexis Drogoul,** who present robot experiments with a method for recording behaviors that will later be reproduced. The approach is a form of explicit teaching and raises the important question of how to generalize learning from the training situation to a later occasion.

In a more cognitively oriented system, **Takashi Omori and Akitoshi Ogawa** apply a model of spatial navigation to the task of symbolic processing, and suggest that attention may be the link between the two. The mechanisms proposed have implications for the understanding of symbolic processing and the emergence of language.

Another developmental architecture is presented by **Georgi Stojanov** who also discusses the relation between the purely behavioristic approach to development and a more cognitive approach. This is likely to become one of the most important questions for epigenetic robotics as it has been in psychology. Although it is often declared that the war between associationists and cognitivists has been won by the latter, the battle does not yet appear to be over.

Hideki Kozima and Hiroyuki Yano describe their ongoing work with a humanoid robot, which in many ways personifies the central ideas, as well as the challenges of the field of Epigenetic Robotics. The *Infanoid* robot is endowed with a "naturalistic embodiment" resembling that of a 3-year-old child. The authors describe how through a combination of "built-in" capacities (a sensori-motor system, reflexes, drives, value-system and learning mechanisms) and physical and social interaction with the environment, a series of cognitive capabilities are hoped to emerge. These include the acquisition of intentionality (goal-directedness), social identification (through joint attention and "action capture"), imitation learning and symbol/language acquisition.

Giorgio Metta, **Giulio Sandini**, **Lorelnzo Natale and Francesco Panerai** describe another humanoid robot, *Babybot,* which is the center of a biologically inspired project that attempts to reproduce a developmental process in a robot. They stress the important roles of action and adaptation in a developing system.

Two papers deal with developmental disorders. **Christian Balkenius and Petra Björne** present a brain-like computational model that is able to reproduce some of the symptoms of ADHD. Although the model does not directly describe the developmental process, it illustrates the important role of context in the inhibition of behavior.

Another study of a developmental disorder is described by **Hei-Rhee Ghim**, **Heyonjin Lee and Sunmi Park** who present an experimental study of autistic children's understanding of false belief (both other's and one's own), showing performance similar to that of children with cognitive deficiencies. Furthermore, it was shown that explicit teaching of the false belief concept had a positive effect in both groups.

The nature of "embodiment"

In an invited contribution, which in a way serves as a "bridge" between the more technical and the more theoretical contributions, **Tom Ziemke** addresses the central concept of *embodiment*, and distinguishes between 5 different notions represented in the literature: (1) *structural coupling,* (2) *physical*, (3*) 'organizmoid'* , (4*) organismic* and (5) *historical embodiment.* The first four form an order of strict inclusion, with organismic embodiment being most restrictive, while (5) characterizes only those systems whose structure is the result of a history of agent-environment interaction. Departing from Searle's well-known "Chinese Room argument" and continuing through the much less-known, but highly relevant work of von Uexküll and Maturana & Varella, the author argues that there is (still) no basis for intrinsic meaning and intentionality in present-day artificial systems of any kind. "Ontogeny preserves the autonomy of an organizations, it does not 'construct' it", states Ziemke, in an open challenge to "emergentist" approaches such as Epigenetic Robotics.

Attention, imitation and the emergence of language

In an important invited contribution **Chris Sinha** clarifies the central concepts of *epigenesis*, *emergence*, *intentionality* and *symbolization* as applied in developmental psychology, all of which are shown to be controversial. The distinction between "signals" and "symbols" is stressed, especially since the second, but not the first require intersubjectivity, which is grounded in joint attention, and fully available only to human beings. Human language is seen as a structural elaboration of symbolization, thus construed. Implicitly, Sinha poses a serious challenge to robotic models, since it is not at all clear how they could be capable of such intersubjective states.

Gedeon Deak, Ian Fasel and Javier Movellan's paper can be seen as a constructive answer to Sinha's challenge. Adopting an adaptive systems approach and relying on recent behavioral data, the authors show inadequacies of two popular models of joint (or shared) attention, and outline a framework for the study of its emergence strongly based on contingency learning and affect. This model appears to be compatible with (epigenetic) robotic modeling. An actual robotic system, implementing some aspects of the theory, is briefly described.

Yuval Marom and Gillian Hayes address a complementary issue: What are the mechanisms of attention as such, and how can attention be implemented in a machine? In the presented model, attention is seen as *dis*habituation, emerging from a dynamic interaction between habituation, novelty detection and forgetting. While applying the model (in a simulated and robotic environment) to non-social tasks such as wall following, the authors nevertheless stress the role of social interaction as a form of stimulus enhancement even for such relatively low-level tasks.

A phenomenon closely related to joint attention in structure and possibly consequences for cognition, which is the focus of much current exploration (e.g. Dautenhahn and Nehaniv, in press) is that of imitation. In their contribution, **Aris Alissandrakis, Chrystopher Nehaniv and Kerstin Dautenhahn** propose a general mechanism for action imitation by agents with different "embodiments", relying on a fairly abstract interpretation of the concept. This mechanism, ALICE, is based on the notion of a *correspondence*, "a recipe through which an imitator can map observed actions of the demonstrator to its own repertoire, constrained by its embodiment and by context". ALICE is illustrated and evaluated in a "chessworld" simulated environment, where the chess figures serve as simulated imitating agents.

In their paper, **Jean Baillie and Chrystopher Nehaniv** use an even more minimal environment of connectionist agents in order to demonstrate an important theoretical point, i.e., under certain conditions, a common "vocabulary" can emerge without an externally defined interaction protocol and criteria of success. In the described simulation, the agents asynchronously "point" to one another and "utter sounds", while only sometimes attending to the "behavior" of the others. A Hebbian learning mechanism nevertheless manages to stabilize a largely shared repertoire of "names".

The minimalist tendency towards embodiment and situatedness observed in the previous two papers is brought to a logical limit in **Anat Treister-Goren and Jason Hutchens's** contribution, describing the development (in the engineering sense of the word) of the "child machine" HAL, which interacts with the outside world only through linguistic "symbols". The authors adopt a strictly behavioristic approach to language learning, endowing their system with two stochastic learning models: one for learning word collocations, and the other based on external reinforcement, administered by HAL's "trainers", who also evaluate the "behavior" of the system relative to human developmental milestones. It is most interesting to look closely at the authors' claim that this approach has lead to "the development of a child machine which can acquire and use language at the level of eighteen-month-old human infant", considering that the guiding assumptions of this work openly contradict those of most other workshop participants.

Kerstin Fischer and Reinhard Morats take a very different approach to human-machine verbal communication, describing a robot capable of language interpretation, spatial orientation and goal-directed motion. In an empirical study, the authors find that users' "mental models" of the robot's abilities determine the way they formulate their linguistic instructions, and when there is a mismatch, the result is "less an effective dialogue with the robot than a depressing or amusing experience of communicative failure". This conclusion highlights the importance for the interlocutor to understand the actual "life-world" of a robot for communication to be possible.

The final two contributions discuss theoretical issues related to the development of language and meaning, applying the discussion to artificial systems. In his paper, **Christopher Prince** argues that "symbol grounding", in the sense of Harnad (1990), underestimates the role of basic human cognitive capacities such as conceptual representation, productivity and systematicity. What is called for, if higher cognitive structures are not to remain purely abstract, is a form of "theory grounding". The author suggests that such grounding should be achieved gradually, throughout ontogenetic development and "artificially intelligent" systems need to recapitulate such development, possibly by combining techniques from 'social robotics' and developmental connectionist learning.

Finally, **Jordan Zlatev** presents the outlines of a multidisciplinary theory of meaning based on the concept of *value* by synthesizing ideas from psychology, semiotics and cybernetics. The author distinguishes between four types of meaning systems: *cue-based, association-based, icon-based,* and *symbol-based*, which he claims form an evolutionary and epigenetic hierarchy. The proposed theory is first briefly applied to evolution, then to ontogeny (showing strong parallels between the two), and finally to the field of developing intelligent artificial autonomous systems. The author's conclusions are pessimistic for existing systems, but not nihilistic for the field of Epigenetic Robotics as such.

Conclusions

This brief review of the papers in this volume shows considerable overlap in the individual approaches, but also substantial differences. We conclude by pointing out three issues that appear to be highly controversial, and therefore we believe they will be the focus of much productive discussion during the workshop:

- What concepts of *embodiment*, *situatedness* and *development* are most adequate for Epigenetic Robotics and for Cognitive Science in general?
- Can present day artificial systems (robotic or not) develop true animal-like and human-like intelligence and autonomy ("Strong AI"), or are they still restricted to being simulations dependent on an external interpretation?

- Is *consciousness* in many of its manifestations: emotions, intentionality, awareness, intersubjectivity etc., an essential requirement for complex cognitive development, and if so, how can artificial systems be made capable of it?

References

Dautenhahn, K. and C. L. Nehaniv (Eds.). (in press) *Imitation in Animals & Artifacts*, Cambridge, Mass.: MIT Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335-346.

Ziemke, T. (Ed.) (in press). Situated and Embodied Cognition, special issue for the *Cognitive Systems Research Journal*.