

## Policy Forum

ARTIFICIAL INTELLIGENCE:

### **Autonomous Mental Development by Robots and Animals**

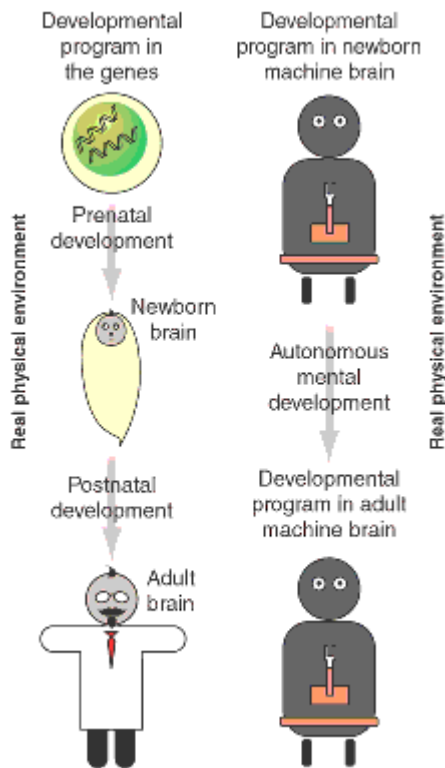
Juyang Weng, <sup>\*</sup> James McClelland, Alex Pentland, Olaf Sporns, Ida Stockman, Mriganka Sur, Esther Thelen

How does one create an intelligent machine? This problem has proven difficult. Over the past several decades, scientists have taken one of three approaches: In the first, which is knowledge-based, an intelligent machine in a laboratory is directly programmed to perform a given task. In a second, learning-based approach, a computer is "spoon-fed" human-edited sensory data while the machine is controlled by a task-specific learning program. Finally, by a "genetic search," robots have evolved through generations by the principle of survival of the fittest, mostly in a computer-simulated virtual world. Although notable, none of these is powerful enough to lead to machines having the complex, diverse, and highly integrated capabilities of an adult brain, such as vision, speech, and language. Nevertheless, these traditional approaches have served as the incubator for the birth and growth of a new direction for machine intelligence: autonomous mental development. As Kuhn wrote ([1](#)), "Failure of existing rules is the prelude to a search for new ones."

#### **A Definition**

What is autonomous mental development? With time, a brainlike natural or an artificial embodied system, under the control of its intrinsic developmental program (coded in the genes or artificially designed) develops mental capabilities through autonomous real-time interactions with its environments (including its own internal environment and components) by using its own sensors and effectors. Traditionally, a machine is not autonomous when it develops its skills, but a human is autonomous throughout its lifelong mental development.

Recent advances in neuroscience illustrate this principle. For example, if the optic nerves originating from the eyes of an animal (i.e., a ferret) are connected into the auditory pathway early in life, the auditory cortex gradually takes on a representation that is normally found in the visual cortex ([2](#)). Further, the "rewired" animals successfully learn to perform vision tasks with the auditory cortex. This discovery suggests that the cortex is governed by developmental principles that work for both visual and auditory signals. In another example,



the developmental program of the monkey brain dynamically selects sensory input, (e.g., three fingers instead of one, as normal), according to the actual sensory signal that is received, and this selection process is active throughout adulthood (3).

Computational modeling of human neural and cognitive development has just started to be a subject of study (4, 5). To be successful, mainstream cognitive psychology needs to advance from explaining psychological phenomena in specific controlled settings toward deriving underlying computational principles of mental development that are applicable to general settings. Such computational studies are necessary for understanding of mind.

The idea of mental development is also applicable to machines, but it has not received serious attention in the artificial intelligence community. In the past, many believed that hand programming alone or task-specific machine learning could be sufficient for constructing an intelligent machine. Nevertheless, recently it was pointed out that to be truly intelligent, machines need autonomous mental development (6). (See the figure, below.)

---

**Growing up.** Mental development is realized through autonomous interactions with the real physical world.

---

### Manual Versus Autonomous Development

The traditional manual development paradigm can be described as follows:

- Start with a task, understood by the human engineer (not the machine).
- Design a task-specific representation.
- Program for the specific task using the representation.
- Run the program on the machine.

If, during program execution, sensory data are used to modify the parameters of the above pre-designed task-specific representation, we say that this is machine learning. In this traditional paradigm, a machine cannot do anything beyond the pre-designed representation. In fact, it does not even "know" what it is doing. All it does is run the program.

The autonomous development paradigm for constructing developmental robots is as follows:

- Design a body according to the robot's ecological working conditions (e.g., on land or under water).
- Design a developmental program.
- At "birth," the robot starts to run the developmental program.
- To develop its mind, humans mentally "raise" the developmental robot by interacting with it in real time.

According to this paradigm, robots should be designed to go through a long period of autonomous mental development, from "infancy" to "adulthood." The essence of mental development is to enable robots to autonomously "live" in the world and to become smart on their own, with some supervision by humans.

Our human genetic program has evolved to use our body well. Analogously, the developmental programs for robots should also be body-specific, or specific to robot "species," as are traditional programs.

However, a developmental program for developing a robot mind must have other properties (see the table) that set it apart from all the traditional programs: It cannot be task-specific, because the tasks are unknown at the time of programming, and the robots should be enabled to do any job that we can teach them. A human can potentially learn to take any job--as a computer scientist, an artist, or a gymnast. The programmer who writes a developmental program for a robot does not know what tasks the future robot owners will be teaching it. Furthermore, a developmental program for robots must be able to generate automatically representations for unknown knowledge and skills. Like humans and animals, the robots must learn in real time while performing "on the fly." A mental developmental process is also an open-ended cumulative process: A robot cannot learn complex skills successfully without first learning necessary simpler skills, e.g., without learning how to hold a pen, the robot will not be able to learn how to write.

<b>DIFFERENCES BETWEEN ROBOT PROGRAMS</b>		
<b>Properties</b>	<b>Traditional</b>	<b>Developmental</b>
Not task specific	No	Yes
Tasks are unknown	No	Yes
Generates a representation	No	Yes

of an unknown task		
Animal-like online learning	No	Yes
Open-ended learning	No	Yes



### Early Prototypes

Early prototypes of developmental robots include Darwin V ([7](#)) and SAIL ([6](#), [8](#), shown below), developed independently around the same time but with very different goals. Darwin V was designed to provide a concrete example for how the computational weights of neural circuits are determined by the behavioral and environmental interactions of an autonomous device. Through real-world interactions with physical objects, Darwin V developed a capability for position-invariant object recognition, allowing a transition from simple behaviors to more complex ones.

CREDIT: J. WENG

The goal of the SAIL developmental robot was to generate automatically representations and architectures for scaling up to more complex capabilities in unconstrained, unknown human environments. For example, after a human pushes the SAIL robot "for a walk" along corridors of a large building, SAIL can navigate on its own in similar environments while "seeing" with its two video cameras. After humans show toys to SAIL and help SAIL's hand to reach

them, SAIL can pay attention to these toys, recognize them, and reach them too. To allow SAIL to learn autonomously, the human robot-sitter lets it explore the world on its own, but encourages and discourages behaviors by pressing its "good" button or "bad" button. Responses invariant to task-unrelated factors are achieved through automatically deriving discriminating features. A real-time speed is reached by self-organizing large memory in a coarse-to-fine way (9). These and other examples that aim at automation of learning [e.g., (10)] have demonstrated robotic capabilities that have not been achieved before or that are very difficult to achieve with traditional methods.

## **The Future**

Computational studies of autonomous mental development should be significantly more tractable than traditional task-specific approaches to constructing intelligent machines and to understanding natural intelligence, because the developmental principles are more general in nature and are simpler than the world around us. For example, the visual world seen by our eyes is very complex. The light that falls on a particular pixel in a camera depends on many factors--lighting, object shape, object surface reflectance, viewing geometry, camera type, and so on. The developmental principles capture major statistical characteristics from visual signals (e.g., the mean and major directions of signal distribution), rather than every aspect of the world that gives rise to these signals. A task-specific programmer, in contrast, must study aspects of the world around the specific task to be learned; this becomes intractable if such a task, such as vision, speech, or language, requires too many diverse capabilities.

This new field will provide a unified framework for many cognitive capabilities--vision, audition, taction, language, planning, decision-making, and task execution. The sharing of common developmental principles by visual and auditory sensing modalities, as recent neuroscience studies have demonstrated, will encourage scientists to further discover underlying developmental principles that are shared, not only by different sensing and effector modalities, but also by different aspects of higher brain functions. Developmental robots can "live" with us and become smarter autonomously, under our human supervision.

It is important for neuroscientists and psychologists to discover computational principles of mental development. And in fact, developmental mechanisms are quantitative in nature at the level of neural cells. The precision of knowledge required to verify these principles on robots will improve our chances of answering some major open questions in cognitive science, such as how the human brain develops a sense of the world around it.

Advances in creating robots capable of autonomous mental development are likely to improve the quality of human life. When robots can autonomously develop capabilities, such as vision, speech, and language, humans will be able to train them using their own communication modes. Developmental robots will

learn to perform dull and repetitive tasks that humans do not like to do, e.g., carrying out missions in demanding environments such as undersea and space exploration and cleaning up nuclear waste.

We believe that there is a need for a special program for funding support of this new field of autonomous mental development. This program should encourage collaboration among fields that study human and machine mental development. Biologically motivated mental development methods for robots and computational modeling of animal mental development should be especially encouraged. There is also a need for a multidisciplinary forum for exchanging the latest research findings in this new field, similar to the Workshop on Development and Learning funded by NSF and Defense Advanced Research Projects Agency held at Michigan State University ([11](#)). We anticipate a potentially large impact on science, society, and the economy by advances in this new direction.

## References and Notes

1. T. S. Kuhn, *The Structure of Scientific Revolution* (Univ. of Chicago Press, Chicago, 3rd ed., 1996), p. 68.
  2. L. von Melchner, S. L. Pallas, M. Sur, *Nature* **404**, 871 (2000).
  3. X. Wang, M. M. Merzenich, K. Sameshima, W. M. Jenkins, *Nature* **378**, 13 (1995).
  4. J. L. Elman *et al.*, *Rethinking Innateness: A Connectionist Perspective on Development* (MIT Press, Cambridge, MA, 1997).
  5. E. Thelen, E. G. Schoner, C. Scheier, L. B. Smith, *Behav. Brain Sci.*, in press.
  6. J. Weng, in *Learning in Computer Vision and Beyond: Development in Visual Communication and Image Processing*, C. W. Chen, Y. Q. Zhang, Eds. (Marcel Dekker, New York, 1998) (Michigan State Univ. tech. rep. CPS 96-60, East Lansing, MI, 1996).
  7. N. Almassy, G. M. Edelman, O. Sporns, *Cereb. Cortex* **8**, 346 (1998).
  8. J. Weng, W. S. Hwang, Y. Zhang, C. Evans, in *Proceedings of the 2nd International Symposium on Humanoid Robots*, 8 to 9 October 1999, Tokyo, pp. 57-64.
  9. W. S. Hwang, J. Weng, *IEEE Trans. Pattern Anal. Machine Intell.* **22**, 11 (2000).
  10. D. Roy, B. Schiele, A. Pentland, in *Workshop on Integrating Speech and Image Understanding, Proceedings of an International Conference on Computer Vision*, 21 September 1999, Corfu, Greece (IEEE Press, New York, 1999).
  11. *Proceedings of Workshop on Development and Learning*, 5 to 7 April 2000, Michigan State University, East Lansing, MI.  
[www.cse.msu.edu/dl/](http://www.cse.msu.edu/dl/).
-

J. Weng is at the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA. J. McClelland is at the Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA. A. Pentland is at The Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. O. Sporns and E. Thelen are at the Department of Psychology, Indiana University, Bloomington, IN 47405, USA. I. Stockman is at the Department of Audiology and Speech Sciences, Michigan State University, East Lansing, MI 48824, USA. M. Sur is at the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

\*To whom correspondence should be addressed. E-mail: [weng@cse.msu.edu](mailto:weng@cse.msu.edu)