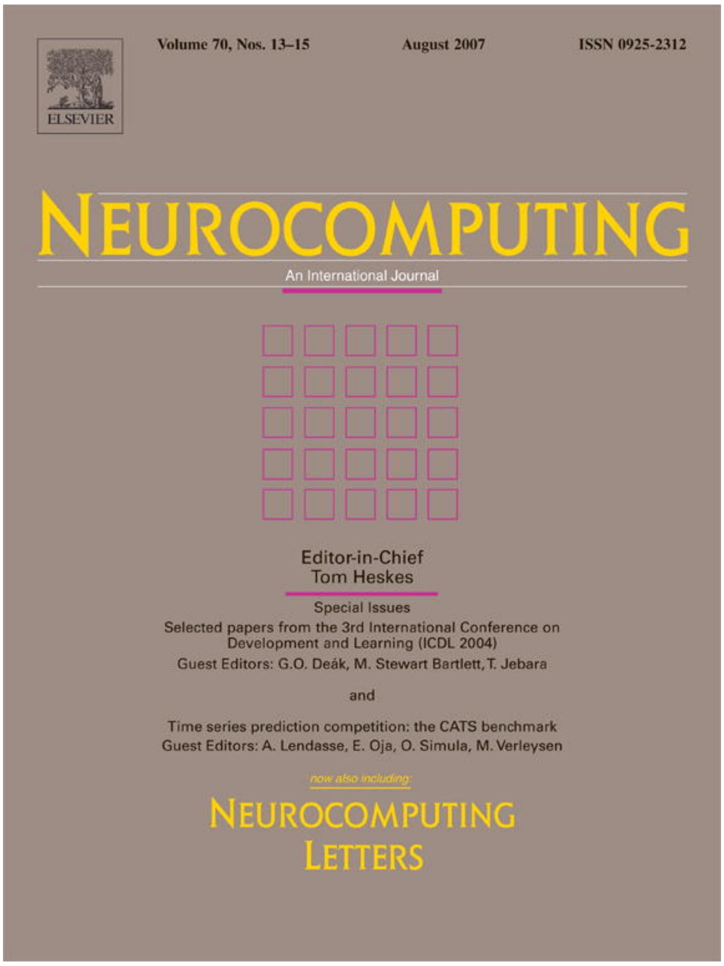
Volume 70, Nos. 13–15
August 2007
ISSN 0925-2312
ELSEVIER
NEUROCOMPUTING
An International Journal
Editor-in-Chief
Tom Heskes
Special Issues
Selected papers from the 3rd International Conference on Development and Learning (ICDL 2004)
Guest Editors: G.O. Deák, M. Stewart Bartlett, T. Jebara
and
Time series prediction competition: the CATS benchmark
Guest Editors: A. Lendasse, E. Oja, O. Simula, M. Verleysen
now also including:
NEUROCOMPUTING LETTERS

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Neurocomputing 70 (2007) 2139–2147
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Editorial

New trends in Cognitive Science: Integrative approaches to learning and development

Gedeon O. Deák[a,*], Marni Stewart Bartlett[b], Tony Jebara[c]

[a]*Department of Cognitive Science, University of California-San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0515, USA*
[b]*Institute for Neural Computation, University of California-San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0523*
[c]*Department of Computer Science, Columbia University, CEPSR rm. 605, 1214 Amsterdam Ave., New York, NY 10027, USA*

Available online 2 January 2007

Abstract

A new trend in Cognitive Science is the use of artificial agents and systems to investigate learning and development of complex organisms in natural environments. This work, in contrast with traditional AI work, takes into account principles of neural development, problems of embodiment, and complexities of the environment. Current and future promises and challenges for this approach are defined and outlined.

Keywords: Active vision; Brain development; Cognitive development; Cognitive Science; Embodied cognition; Emergent properties; Ethnography; Interdisciplinary research; Language learning; Learning algorithms; Machine learning; Modeling; Social communication

1. Historical changes in modeling: towards adaptive, interactive intelligence

A widely circulated story about early Cognitive Science efforts has it that Marvin Minsky assigned an undergraduate student "computer vision" as a summer project. The anecdote builds irony with historical hindsight: The optimism of early AI work, like successive theoretical trends in the behavioral and computational sciences, ran into the barrier of Real Complexity: monumentally interactive intricacies of organism–environment dynamics that give rise to human thought and action. Cognitive Scientists have repeatedly discovered that prosaic skills like producing phonemes or tracking objects are quite challenging to capture in models that approach the scale of an real organism. A growing appreciation of behavioral and cognitive details the complexity of anatomical structure and function of real brains and bodies, and the difficulties of describing ecological structures jointly mandate a reconceptualization of models of intelligent behavior. The mandate is to push beyond the symbolic models of human information processing of the 1980s and 1990s, and to meaningfully elaborate on early work on neural networks by incorporating relevant information about neuroscience (e.g., chemistry, physiology, anatomy), concerns about embodiment (e.g., perception–action systems, biomechanics, motor control), and sensitivity to cognitive ecology (e.g., ethnographic data at different levels of detail). Of course, such models must also accurately model high-quality behavioral data from organisms of interest, be they rat or human, infant or adult.

In recent years a community of researchers has made strides in these theoretical and empirical directions. Their work is bringing new questions and problems to the foreground, and demonstrating innovative empirical approaches, as exemplified by contributions to this issue. Although the methods and questions are quite varied, a "family resemblance" of recurring concerns or positions can be discerned (though perhaps not all the contributors would agree with all these formulations):

- Cognition occurs in the context of complex structures, both physical and social, in the environment. In many ways this structure alters, and even sometimes simplifies,

*Corresponding author. Tel.: +1 858 382 7306; fax: +1 858 534 1128.
E-mail addresses: deak@cogsci.ucsd.edu (G.O. Deák), marni@salk.edu (M.S. Bartlett), jebara@cs.columbia.edu (T. Jebara).

0925-2312/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2006.06.008

the computational and behavioral tasks faced by active agents in those environments.
- Problems of embodiment are substantial and important [13]. We cannot fully understand intelligent behavior without understanding how the information-processing system is integrated with, and shaped by, its ''platform'' for acting and perceiving.
- Cognition is sub-symbolic and distributed. However, it is not all of one simple type (e.g., unsupervised Hebbian learning). In discerning types of learning and cognition, we require contributions from theoretical and experimental neuroscience.
- Nativist accounts that directly ascribe complex cognition to experience-independent products of the genome are not consistent with developmental neuroscience, embryology, or genetics, or with general principles of epigenesis. Such accounts, while sometimes pragmatic in initial models of cognition, are unparsimonious and should be marked as simplified and speculative.

More generally, many proponents of these positions and others have questioned or even rejected the traditional theoretical framework of cognitive psychology and AI. As summarized by Christiansen and Hooker [12], most theories in cognitive science implicitly assume general *centralized control models*. Such models place a disembodied, Cartesian mind at the center of a Ptolemean cognitive universe, wherein the environment (including the body) is separate and subordinate [17,44]. This standard model ignores issues of embodiment and the environment. This is theoretically problematic [12], and contrary to findings from many disciplines. For example, Pentland [43] has shown that a great deal of people's impending behavior can be predicted by where they are and who they are with. Note that not all the social sciences have historically subscribed to this centralized control model: for example, the opposite problem can be seen in pure ethnographic approaches that emphasize relativism, where the environment is given full causal power without considering shared neural and perceptual–motor attributes of individuals within and between cultural groups. The alternative is to reject ''either/or'' models of both extremes. Instead, we assume that adequate models of cognitive functions require an accurate account of the tendencies and variability of real behavior, a detailed model of the body that provides for and executes the brain's computation, a detailed model of the functions and processes of the neural systems, and detailed models and descriptions of patterns of information on various spatial, temporal, and cultural scales, within the environment. Put otherwise, we assume the major challenge for Cognitive Science is Systems Modeling.

Systems modeling: Testing and falsifying formal theories about specific cognitive functions of organisms with vastly complex nervous system and vastly complex perceptual and motor capacities, interacting in real time and space with highly diverse and changeable environments.

Some proponents go further in breaking from traditional AI, cognitive psychology, linguistics, and anthropology/sociology. Historically, these fields have mostly ignored developmental/epigenetic concerns. Now, however, we know enough about brain development, and about socio-cultural effects on infants'/children's thinking, to infer that a developmental history must be part of any account of mature cognitive functioning. A description of the mature cognitive ''profile,'' while necessary, cannot yield a full explanatory account. A full explanation must include an account of how that profile emerged. A non-developmental view of adult cognition can lead to systematic misconceptions about the adult profile [36,55]. Mature functioning is a product of protracted learning and development in constant interaction with genetically mediated, heterochronous processes of neural change [22,32,46]. Thus, although most papers in this volume do not explicitly consider children of a particular age, or seek to model precise developmental changes, many deal with cognitive processes that are centrally relevant to infants and children, and are controversial: for example, face processing [Bartlett], inductive inference [Nelson and Cottrell], and syntax acquisition [Desai]. Such difficult natural learning problems can be re-cast in developmental terms. This in turn calls for developmental systems modeling.

Developmental systems modeling: We define this kind of modeling as follows: Testing and falsifying formal theories about specific cognitive functions of *developing* organisms with *emerging,* vastly complex nervous system and *emerging* vastly complex afferent and efferent potentials, *through a history of interaction* in real time and space within highly diverse environments that change *on multiple time scales, ranging from moment-by-moment changes to long-term changes over the organism's lifespan*.

Many proponents of modern approaches to developmental systems modeling realize the limits imposed by disciplinary boundaries, and seek inspiration from multiple disciplines. For example, computational models often can be improved by careful attention to what is neurally plausible, to are the precise details of human behavior and cognition. Psychology research benefits from richer grounding in the neural underpinnings of thought and behavior, and from rigorous, well-specified process models of cognition. Cognitive neuroscience benefits from a deeper grasp of how organisms perceive and act in natural environments. All of these disciplines can benefit from greater knowledge of biophysics, embryology, ethnography, genetics, linguistics, physical anthropology, and animal behavior.

What empirical problems are of interest to modern proponents of developmental systems modeling? The wide-ranging list is challenging, controversial, and substantive. It includes such problems as face processing, scene processing, attention, word learning, imitation, shared attention, working memory, navigation, articulation, multimodal perception, fluid motor control, self-awareness, object recognition, and others. The papers in this issue exemplify only a few possible interdisciplinary approaches to a few trenchant problems in cognitive science. They were borne of an effort to gain insights about autonomous learning and development by creating a forum for researchers in machine learning, robotics, neuroscience, and developmental psychology.

The International Conferences on Development and Learning (ICDL) have a brief but energetic history. They began as a Workshop on Development and Learning funded by NSF and DARPA in 2000. The next meeting was hosted in 2002 by MIT; the third by UC-San Diego and the Salk Institute in 2004; the fourth by Osaka University in 2005; and the fifth by Indiana University at Bloomington in 2006. The next ICDL meeting will be held in London, UK, in 2007. The papers in this issue were submitted for peer review as expanded versions of presentations from the 2004 meeting.

We now consider how the contributions to this volume exemplify and advance interdisciplinary approaches to modeling developing complex organisms and agents. We organize the discussion around the theoretical challenges addressed by the papers. We close by briefly considering future directions heralded by these contributions.

2. Current challenges in developmental systems modeling

2.1. Challenge #1: Modeling the environment

2.1.1. Overview

A major challenge in effective theory-building in cognitive science is deriving rich and accurate descriptions of agents' environments. Traditional AI methods of representing the environment as highly reduced binary input vectors, or, worse, symbolic abstractions, are limiting. Symbolic approaches tell us little about how (developing) brains learn about environments, or how the structure of the body, the physical environment, or the social environment constrain the agent's learning. By contrast, connectionist models that use sub-symbolic input vectors avoid some of these limitations, but still do not capture the precise ecological structure of the sensorimotor and ecological information that drive information processing. In other words, modern modeling efforts faces a *dual modeling problem*: first, deriving good models of neural learning; second, deriving accurate models of the environment. If one is testing, say, models of statistical learning, the plausibility of the results (in terms of "fit" to a real organism's learning) are a function of the accuracy of the modeled cognition *and* the modeled to-be-learned information.

A major problem in modeling to-be-learned information is that the environment, even if it is greatly reduced, contains a lot of information. This is true even if we consider information in, say, two spectra (e.g., visible light and audible speech sound), from a limited sampling source (e.g., one camera and microphone) in a single setting (e.g., driver's seat of a car). To get a sense of how much information there is, consider that the computer gaming industry spends millions of dollars and the best computer graphics techniques to develop "realistic" simulated environments. Yet the best results, if pleasing, are still only highly reduced simulations of two spectra (visible light in a highly quantized two-dimensional (2D) field; audible sound in a reduced frequency range), in two dimensions in a limited field. The results could never be mistaken for a "real" environment, and permit very, very limited embodied interaction.

2.1.2. Physical structure

In regard to analysis of the physical environment, computer vision has come further than other domains. Machine perception researchers generally care about systems that accurately, rapidly analyze high-quality rich images or (preferably) video. (By contrast, for example, *most* linguists do not test theories with high-fidelity audio recordings of natural speech, much less multi-modal contextual information about the social context of utterances.) Developmental psychology is in a primitive state as well, with little expectation that theories must incorporate details of infants' and children's environments (despite some instructive examples [50]). For example, after decades of vigorous laboratory work there is still *no* data on young infants' everyday looking behaviors. This complacency has had a major impact on theory. For example, neonativist theories of the 1990s [2,56] were based upon stripped-down laboratory studies, with no converging evidence from naturalistic behavior, or analysis of the information patterns available to babies in the first weeks or months of life. It has now been shown that a simple learning agent can develop social categories (e.g., faces) from natural environments after as little as a few minutes of exposure [11]. Thus, neonativist accounts of infants' looking behaviors are unparsimonious, and they bear the burden of proof that specialized learning processes are congenitally available for high-level visual tasks (e.g., counting objects; reasoning about occlusion).

Interest in the structure of information in real environments, and learnability of that information, is evident in the papers that follow. Oh and Choe (this issue) illustrate the importance of self-produced motion by demonstrating that simple neural networks learn texture segmentation better when available information is complex three-dimensional (3D) input, rather than intermediate 2D images. Because 3D input is provided by self-motion, this work illustrates the power of simultaneously modeling ecological information patterns and aspects of embodiment. This supports Gibson's [26] and Ballard's [4] idea that many difficult perceptual problems are simplified when the visual agent moves in the environment. It also shows that isolating visual functions with the simplest possible stimuli is not necessarily the best approach to understanding natural vision.

2.1.3. Social structure

The social structure of the environment is especially complex: people and organisms are highly structured as objects, and they dynamically change in hard-to-describe and hard-to-predict ways. We are all experts in faces, for example, yet human faces are dauntingly complex, as machine learning studies have shown. Bartlett (this issue) explores models of optimal neural coding (i.e., information maximization) in face processing as a paradigm of high-level visual processing. A large body of work has shown relationships between the statistics of the environment and neural coding in early vision [54]. These principles relate to higher visual functions including face recognition. Recent work shows how cognitive phenomena in face processing such as typicality effects, other-race effects, and face adaptation can emerge from a system optimized to the statistics of real face images. In addition, computer face recognition studies show that more successful algorithms for complex perceptual tasks like face processing are adapted to the complex structure of information in the visual environment (i.e., 2D projections of real faces). This line of research implies that the human visual system has developed neural computations optimized to the probability structure of the visual environment, and insofar as face recognition is a crucial social function, the visual system seems to have learned the probability structure of an aspect of the social environment as well. Other work [25] suggests that optimization is not innately specialized for faces, but is plastic during development and even adulthood. A critical future direction will be to improve contextualized face processing systems: robots that move themselves through complex environments and derive information about multiple, unpredictable, 3D dynamic faces.

There are of course many other daunting problems of the social environment, and we cannot easily intuit how much social-cognitive "work" is facilitated by patterns (or frustrated by noise!) in social environments. Cognitive ethnography work has led many scholars to conclude that the individual brain should not be the only cognitive unit of analysis [30]. Historically, though, these insights have not always guided theories of development. In child language, for example, Chomskian theorists assumed that syntactic structures are unlearnable, and therefore innate [15]. This was not based on any analysis of the language infants actually hear. In fact, when the matter has been tested, sufficient information has been found in parental speech to support the gradual acquisition of seemingly "unlearnable" syntactic constraints [38] (see also Refs. [45,47]).

Developmentalists also have posited high-level innate skills that are far removed from either neuroscience data, or rigorous formalization of the information processing necessary to carry out the skill in question. For example, claims about neonatal imitation [40] have been challenged by more parsimonious and plausible accounts [1,33,34]. Imitation seems to be a learned skill constructed from more basic behaviors in complex social environments. This is supported by Zukow-Goldring and Arbib's (this issue) evidence that manual imitation emerges through social input from caregivers. Parents scaffold infants' interactions with objects to facilitate their discovery of objects' affordances. Infants do not acquire object-using skills from observation, but (literally) from hands-on interactions wherein caregivers manually help infants utilize tools. Without documenting such interactions we might blindly attribute infants' imitation, not to mention tool-using skills, to innate capacities. The same paper raises questions about how much infants can make mental-state and causal inferences [59], and how much these inferences are scaffolded by social experience. It also raises intriguing questions about what kinds of human learning interactions should be considered supervised, unsupervised, or semi-supervised (see Section 2.2). For example, when adults hold babies' hands and jointly manipulate an object, what kind of input is this? What do infants learn from such input, compared to just observing an adult doing the actions? Without ethnographic data, we would not recognize the depth of these questions for framing behavioral and computational research questions.

Another innovative interdisciplinary approach to modeling the environment is seen in Yu, Ballard and Aslin's (this issue) contribution. They ask what information caregivers might provide to help pre-linguistic infants acquire word-to-world mappings: a venerable philosophical problem. Yu et al. test the roles of speaker's gaze-cues (i.e., fixations on named referents, or elsewhere [3]), and prosodic cues [23] as two sources of social information that might help infants disambiguate meaning. The researchers used eye tracking and acoustic analysis to capture the structure of these behaviors while people read picture books, and then carried out simulations to test the learnability of this structure. The results show that there is sufficient information in interactions such as picture-book reading to support inferences about what the reader is naming. This addresses debates about the nature of word learning in children [18,39]. It also suggest powerful studies of social exchanges which move from speculation to testing what cues we provide to one another, and to infants and children, when interacting with them.

2.1.4. Interdisciplinary experiments on social response

One approach to studying social-cognitive development is to utilize controlled social agents as stimuli. Some researchers, for example, use robots to study people's generalized responses to well-controlled (and simplified) social agents [9,41]. Oberman, McCleery, Ramachandran and Piñeda (this issue) show that robotic actions, even those without objects, can activate the human mirror neuron system. They show that a neural signal associated with the mirror system (sensorimotor cortex suppression of *mu*-band EEG oscillation) is activated in adults by the sight of a robotic arm making an anthropomorphic grasp, similar to a human arm. This opens up new possibilities for studying the physical stimulus properties that cause the human brain to respond to something as another social being.

A different experimental approach is taken by Teuscher and Triesch (this issue). They test a model of the emergence of gaze-following: a gradually-emerging attention-sharing skill in human infants that is associated with later social learning, including language development [60]. The authors vary properties of the teaching signal, notably the predictability of a simulated "caregiver" that produces sequences of actions from which a simulated infant might notice, and learn, associations between the caregiver's head pose and possible locations of interesting things in the surrounding environment. The results support specific predictions [60] that changes in the caregiver's behaviors (e.g., "personality") will have systematic effects on infants' social skill development. This is a rare developmental model of the emergence of social skills that takes into account neural reinforcement [58] and contingency learning [62], as well as infants' learning capacities (e.g., habituation) and stimulus preferences. Although the authors do not model precise behavioral data, the effects of caregiver interaction style on infant social development has been documented [14], and is being studied in more detail [19].

2.2. Challenge #2: Computational models of learning

2.2.1. Machine learning approaches to theories of learning and development

An important principle when studying the behavior of agents and systems in complex environments is grounding models in real sensor data. This encourages a bottom–up approach to modeling where models are estimated primarily to fit empirical observations instead of relying heavily on a priori theoretical assumptions or expert knowledge [24]. Hypotheses need not be conjectured a priori and tested and rejected through controlled experimentation. In a data-driven approach, hypotheses and models can emerge a posteriori from empirical data on behavior and activity. The branch of artificial intelligence that focuses on data-driven modeling is known as machine learning and has enjoyed considerable advances in the past decade [37,48,57]. Machine learning lies at the intersection of many fields including statistics, computer science, neuroscience, physics, cognitive science, mathematics, and operations research. It deals primarily with data-driven statistical, probabilistic, and computational models. As rich data sources (neural, behavioral, and ecological/ethnographic) become available, it is increasingly advantageous to use models that automatically exploit data instead of relying on manual expert knowledge. This is particularly true for visual and behavioral data, which tend to be not only complex and multidimensional, but also stochastic, approximate and incomplete.

2.2.2. Unsupervised approaches

While an overview of machine learning is beyond the scope of this introduction, several tools have proven useful in modeling neural and behavioral activity (both developmental and non-developmental), and have generated biologically plausible, data-driven models. One particularly useful split of the field is into the categories of unsupervised and supervised learning. Finer splits can be made to identify hybrid categories such as semi-supervised learning and reinforcement learning which lie in the intersection between supervised and unsupervised approaches. In unsupervised approaches, general principles are used to uncover structure from data. For instance, information theoretic criteria including maximum mutual information, maximum entropy, and minimum relative entropy can be exploited as shown by Bartlett (this issue) and Bell and Sejnowski [5] for visual representations. Many variants of unsupervised learning are closely interrelated. For example, an equivalence between information theoretic maximum entropy and the maximum likelihood approach [6] is well known.

Furthermore, maximum likelihood itself can be viewed as a variant of Bayesian learning where a single point estimate is used as a surrogate for a distribution over all possible models. By exploiting Bayes' rule and considering distributions over models and hypotheses, a fully Bayesian approach to inference is utilized by Nelson and Cottrell (this volume) for problems of object categorization and concept learning. The Bayesian framework and Bayes' rule are not limited to pairs of variables, but can accommodate highly structured multivariate networks. Recent efforts have married principles of Bayesian inference with graph theory [42] to offer a principled way of performing inference on large-scale multivariate problems. As the number of variables grows, Bayesian methods also offer a natural way of controlling model complexity [28], although other approaches such as minimum description length [49] are also viable. Notably, there is now evidence that some neural processes can be described in terms of Bayesian inference [63]. The limits of this approach continue to be explored by Nelson and others.

2.2.3. Supervised and semi-supervised approaches

On the other extreme, supervised learning methods exist which discriminatively focus on the learning problem of predicting outputs from inputs. Often, however, supervised data collection is more expensive since an expert teacher needs to label or identify correct outputs for each input. Model complexity is once again kept in check in supervised approaches. The methods of choice for complexity control for supervised learning include regularization principles [27] or statistical generalization guarantees such as structural risk minimization and the Vapnik–Chervonenkis dimension [61]. Such supervised methods have been very successful in applied domains. For instance, supervised learning has been used to uncover mappings from neuron activation levels in the macaque cortex to visual object category [29]. A weaker form of supervision, which is less cumbersome for data collection process, is reinforcement learning. Here, the teacher only provides a binary feedback if a system produces the correct output or action given the input [35]. A reinforcement learning approach is exploited by Teuscher and Triesh (this issue) in order to learn to predict locations of rewards (conceived as interesting objects from a simulated caregiver). This approach not only models some typical developmental sequences but also some differences in developmental disorders by making reasonable changes to parameters of the model. Again, this machine learning approach is supported by neuroscience data (see Section 2.3), and holds promise for developmental theories [52].

Computational techniques more familiar to psychologists, such as variants of back-propogation neural networks, offer viable approaches to many problems of learning and development [53]. One application is to neural processes of sensory coding and perceptual classification. Yoshikawa, Hosoda and Asada (this issue) address a famous problem of self-perception: how does the infant learn to differentiate self from other? This problem was addressed in a philosophical vein by scholars like Freud and Piaget, but the authors show it to be amenable to computational formalisms. They use an elegant ''cross-anchoring'' Hebbian learning procedure to demonstrate how the infant might learn to differentiate self-touching events from touches by others. It does so by learning cross-modal associations between regions where a touch is felt and simultaneously a visual occlusion (as by a hand touching part of a leg) is seen. This initial work, also implemented in a simple robot, lays the groundwork for sophisticated studies of the emergence of intermodal knowledge in. Yet sensorimotor development is not the sole domain of computational simulations. Models of comparable complexity have been used to simulate high-level language development [7,21]. Desai (this issue), in this tradition, models language input patterns akin to certain syntactic structures, and attempts to simulate a curious developmental phenomenon: the shift with age from using general context to word-specific syntax and semantics to correctly use verbs. For example, an English-speaking toddler's production *''Don't fall that on me!'' [8], while sensible, assigns non-standard causativity (and transitivity) to the verb *fall*. With development, children converge on lexically specific patterns of use. Some researchers explain this shift as a competition between innate biases and specific lexical knowledge. Desai shows that a recurrent network can learn to predict constituent order in sentences that might be either transitive or intransitive. The results suggest a shift from learning general sentence-level cues to using verb-specific patterns as experience with specific verbs accrues. Thus, language-specific innate principles are not necessary to explain the shift from frame-to verb compliance.

2.2.4. Continuing debates; future directions

The artificial intelligence and learning community has frequently debated the tradeoff between a priori model structures (i.e., specifying expert domain knowledge) and how much should be determined by data (i.e., a black box learner). Early AI hand-designed symbolic approaches were challenged by generic tools like the perceptron, but regained ground due to the linearity limitations of the perceptron. Neural networks won back ground for black-box approaches by extending the flexibility of perceptrons via nonlinear settings. The arrival of Bayesian networks [42] in the 1990s pushed the pendulum back towards more expert-driven methods, but a resurgence of black-box methods is at hand with supervised support vector machines [61] and regularization theory [27]. In fact, these approaches can be extended to accommodate structured knowledge about the problem domain, which further improves performance [31]. Thus, the tradeoff between domain knowledge and generic flexibility still exists, and that is reflected in this issue, as in the contrast between Nelson and Cottrell's approach and Desai's.

Another open issue in machine learning is how to represent data so models are more reliably applicable and generalizable. Many measurements from sensory, behavioral, neural, and environmental data sets are not just simple vectors or discrete states. How can such data be represented computationally to be more compatible with our learning algorithms? How can more flexible representations and more powerful algorithms improve learning without requiring unrealistically protracted training or feedback conditions? How can we incorporate a priori knowledge and partial information without over-constraining models? Ultimately, machine learning is a versatile tool if we do not limit it to pure black boxes, or resort to symbolic approaches. More bridges to embodied developmental, neural, and ecological data are needed for the impact of machine learning work to be realized in the cognitive and developmental sciences.

2.3. Challenge #3: Developmental systems modeling

2.3.1. Brain, body, experience, and development

As interdisciplinary attempts to test theories of learning and development have gained sophistication, formal models and simulations have been changing. This is seen in the interface of machine learning and developmental psychology. Research at this border zone has sharpened questions of embodiment [13]. *Embodiment* is the buzzword for a break with radical functionalism of circa-1980s AI (i.e., the idea that implementation of a cognitive process is irrelevant, and various abstractions, manifested in neural or silicon systems, can be treated as theoretically equivalent). A central idea is that cognition and behavior cannot be understood just as cognitive-symbolic abstractions, but as physical functions for control of real bodies that interact with a real world. These functions determine not just how the brain functions, but how it develops in an organism's history. Some controversial topics in human development centrally involve questions of embodiment: for example, imitation (Section 2.1.3) [10,20,51]. This is an good case study for embodiment because we use motor actuators to imitate other agents with similar motor capacities. Infants must somehow learn mappings between another's actions and their own—a process no less challenging to understand than to implement.

2.3.2. Developmental systems theories

In developmental psychology, Karmiloff-Smith [36] and Elman et al. [22] outlined a "neuro-constructivist approach" to explaining the autonomous development of cognition and behavior in children. Two contributions to this issue outline formalizations and approaches for explaining autonomous cognitive development. Dominey's (this issue) constructivist model takes perceptual primitives (events and objects) from simple physical environments, learns to map them to symbols, and learns to represent strings of symbols in syntactic constructions that support acquisition of abstract semantic–syntactic structures (e.g., sentence roles). Dominey argues that the same cognitive architecture that learns event abstractions can also learn syntactic constructions, and, perhaps, social categories. This model departs markedly from models that assume modularity of syntactic knowledge, even with respect to other high-level symbolic and representational knowledge. Importantly, it addresses the difficult question of how syntax is related to embodied experience. It will be intriguing to see this model can be expanded to predict behavioral and neurological data from a broad range of domains.

Weng (this issue) describes embodied implementations of machine learning, also encompassing a range of behaviors, from drawing to navigation. His paper provides a valuable conceptual framework for understanding what is meant by "autonomous development," and for classifying and comparing computational and robotic approaches ranging from disembodied toy models, to minimally embodied virtual agents or simple robots, to robots that can generate genuinely new behaviors. Importantly, he ventures to address the difficult problem of how autonomous agents could acquire self-knowledge and use this knowledge to drive adaptive learning. Metacognitive development, and its role in cognitive development and skill acquisition, is not well understood [16], but because of its clear importance for learning and self-regulation, it will be intriguing to see whether future robotic approaches yield theoretical insights into the nature of self-aware learning.

3. Conclusions

This overview and the papers in this issue address only a small portion of the potential research questions on modeling developmental systems. There are many challenges on the horizon. One challenge is extending current learning theories to high-level social, linguistic, and conceptual knowledge. It is difficult enough to demonstrate the viability of formal models of learning to follow gaze, or recognize faces. It is another to generate viable theories of how children learn to lie, or understand that different people have different beliefs or preferences, or learn to joke around with other people. An equally difficult problem is to understand how any of these developmental outcomes are implemented in real, developing neural systems and processes.

Another challenge concerns the use of various simulation platforms (digital, virtual, robotic) to test different models of learning and development. Regardless of platform, the kinds of formal learning algorithms used and their relation to neural systems must be considered. However, the current interdisciplinary community lacks an agreed-upon heuristic for matching the "right" platform to the right problem. Sometimes, as a result, it is not clear why a robot has been used or what has been learned as a result.

A third challenge is modeling the environment. Simplified, disembodied input strings might be appropriate for testing some models, but in most cases we cannot understand learning and development without a rich, theoretically informed descriptions of the environment. Dense ethnographic datasets (including audio and video records) are necessary for deriving good descriptions of the ecological information structures available to learners. However, good ethnographic studies present their own challenges, not least of which is that, using traditional methods, they are extremely laborious and slow.

A fourth challenge concerns development. Both bodies and brains change with development, and partly as an effect of experience. With these changes come shifting experiences of the environment. Much of this change is non-monotonous, non-linear, and heterochronous. This means our models of development cannot be overly simplistic. Collaborations between developmental researchers, machine learning researchers or roboticists, and neuroscientists will be necessary to make real progress on complex, well-specified models of development.

In sum, this special issue is a sampling of innovative efforts to address challenging issues on development and learning in embodied systems operating in natural environments. These issues involve complex questions of brain–behavior, brain–body, and organism–environment interactions throughout development. For this, we need the knowledge the methods, and the theory-building tools from multiple disciplines. The International Conferences on Learning and Development support a growing, vibrant community of researchers who share this goal.

References

[1] M. Anisfeld, Only tongue protrusion is matched by neonates, Dev. Rev. 16 (1996) 149–161.

[2] R. Baillargeon, The object concept revisited: New directions in the investigation of infants' physical knowledge, in: C. Granrud (Ed.), Visual Perception and Cognition in Infancy, Erlbaum, Hillsdale, NJ, 1993, pp. 265– 315.

[3] D.A. Baldwin, Understanding the link between joint attention and language, in: C. Moore, P. Dunham (Eds.), Joint Attention: Its Origins and Role in Development, Erlbaum, Hillsdale, NJ, 1995.

[4] D. Ballard, Principles of animate vision, Artif. Intell. 48 (1991) 57–86.

[5] A.J. Bell, T.J. Sejnowski, An information maximization approach to blind separation and blind deconvolution, Neural Comput. 7 (6) (1995) 1129–1159.
[6] A. Berger, S. Della Pietra, V. Della Pietra, A maximum entropy approach to natural language processing, Comput Linguistics 22 (1996).
[7] A. Borovsky, J.L. Elman, Language input and semantic categories: a relation between cognition and early word learning, J. Child Lang., in press.
[8] M. Bowerman, Starting to talk worse: clues to language acquisition from children's late speech errors, in: S. Strauss (Ed.), U-shaped Behavioral Growth, Academic Press, New York, 1982.
[9] C. Breazeal, Emotion and sociable humanoid robots, Int. J. Hum.–Comput. Stud. 59 (2003) 119–155.
[10] C. Breazeal, B. Scassellati, Robots that imitate humans, Trends Cogn. Sci. 8 (2002) 481–487.
[11] N. Butko, I. Fasel, J. Movellan, Learning about humans during the first 6 minutes of life, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[12] W.D. Christensen, C.A. Hooker, An interactivist-constructivist approach to intelligence: self-directed anticipative learning, Philos. Psychol. 13 (2000) 5–45.
[13] A. Clark, Being There: Putting Brain, Body, and World Together Again, MIT Press, Cambridge, MA, 1997.
[14] F. Cohn, R. Matias, E.Z. Tronick, D. Connell, D. Lyons-Ruth, Face-to-face interactions of depressed mothers and their infants, in: E. Tronick, T. Field (Eds.), Maternal Depression and Infant Disturbance, Jossey-Bass, San Francisco, 1986, pp. 31–45.
[15] S. Crain, 1991 Language acquisition in the absence of experience, Brain Behav. 14 (1991) 597–611.
[16] K. Crowley, J. Shrager, R.S. Siegler, Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms, Dev. Rev. 17 (1997) 462–489.
[17] A.R. Damasio, Descartes' Error: Emotion, Reason, and the Human Brain, Putnam, New York, 1994).
[18] G.O. Deák, Chasing the fox of word learning: why "constraints" fail to capture it, Dev. Rev. 20 (2000) 29–80.
[19] G. Deák, A. Krasno, H. Jasso, Y. You, Watch the hands: infants learn gaze-following from parents' pointing and manual actions, Presented at the XVth Biennial International Conference on Infant Studies, Kyoto, Japan, June 2006.
[20] Y. Demiris, A. Dearden, From motor babbling to hierarchical learning by imitation: a robot developmental pathway, in: Proceedings of the Fifth International Workshop on Epigenetic Robotics, Lund University, Cogn. Stud. 123 (2005) 31–37.
[21] J.L. Elman, Connectionist models of cognitive development: where next?, Trends Cogn. Sci. 9 (2005) 111–117.
[22] J. Elman, E. Bates, M.H. Johnson, A. Karmiloff-Smith, D. Parisi, K. Plunkett, Rethinking Innateness: A Connectionist Perspective on Development, MIT Press, Cambridge, MA, 1996.
[23] A. Fernald, C. Mazzie, Prosody and focus in speech to infants and adults, Dev. Psychol. 27 (1991) 209–221.
[24] T. Fumihide, J. R. Movellan, The RUBI Project: designing everyday robots by immersion, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[25] I. Gauthier, M.J. Tarr, A.W. Anderson, P. Skudlarksi, J.C. Gore, Activation of the middle fusiform face area increases with expertise in recognizing novel objects, Nat. Neurosci. 2 (1999) 568–573.
[26] J.J. Gibson, An ecological approach to visual perception (1979).
[27] F. Girosi, M. Jones, T. Poggio, Regularization theory and neural network architectures, Neural Comput. 7 (1995) 219–269.
[28] D. Heckerman, A tutorial on learning with Bayesian networks, in: M. Jordan (Ed.), Learning in Graphical Models, MIT Press, Cambridge, MA, 1999.
[29] C.P. Hung, G. Kreiman, T. Poggio, J.J. DiCarlo, Fast readout of object identity from Macaque inferior temporal cortex, Science 310 (2005).
[30] E. Hutchins, Cognition in the Wild, MIT, Cambridge, MA, 1995.
[31] T. Jebara, Machine Learning: Discriminative and Generative, Kluwer Academic, Dordrecht, 2004.
[32] M.K. Johnson, Developmental Cognitive Neuroscience, Blackwell, Hove, UK, 1997.
[33] S.S. Jones, Imitation or exploration, Child Dev. 67 (1996) 1952–1969.
[34] S.S. Jones, Infants learn to imitate by being imitated, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[35] L.P. Kaelbling, M.L. Littman, Reinforcement learning: a survey, J. Artif. Intell. Res. 4 (1996).
[36] A. Karmilloff-Smith, Development itself is the key to understanding developmental disorders, Trends Cogn. Sci. 2 (1998) 389–398.
[37] T.S. Lee, S.X. Yu, An information-theoretic framework for understanding saccadic eye movements, in: S.A. Solla, T.K. Leen, K.-R. Müller (Eds.), Advances in Neural Information Processing Systems, vol. 12, MIT Press, Cambridge, MA, 2000, pp. 834–840.
[38] J.D. Lewis, J.L. Elman, A connectionist investigation of linguistic arguments from poverty of the stimulus: learning the unlearnable, in: Proceedings of the 23rd Annual Conference of the Cognitive Science Society, 2001, pp. 552–557.
[39] E.M. Markman, Constraints on word meaning in early language acquisition, Lingua 92 (1994) 199–227.
[40] A.N. Meltzoff, M.K. Moore, Persons and representation: why infant imitation is important for theories of human development, in: J. Nadel, G. Butterworth (Eds.), Imitation in Infancy, CUP, Cambridge, UK, 1999, pp. 9–35.
[41] J.R. Movellan, J.S. Watson, The development of gaze following as a Bayesian systems identification problem, Int. Conf. Dev. Learning 2 (2002) 34–40.
[42] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, 1997.
[43] A. Pentland, Are we one? On the nature of human intelligence, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[44] N. Praetorius, Inconsistencies in the assumptions of constructivism and naturalism: an alternative view, Theory Psychol. 13 (2003) 511–539.
[45] G.K. Pullum, B.C. Scholz, Empirical assessment of stimulus poverty arguments, Linguistic Rev. 19 (2002) 9–50.
[46] S.R. Quartz, T.J. Sejnowski, The neural basis of cognitive development: a constructivist manifesto, Behav. Brain Sci. 20 (1997) 537–596.
[47] T. Regier, S. Gahl, Learning the unlearnable: the role of missing evidence, Cognition 93 (2004) 147–155.
[48] L. Renninger, Walker, J. Coughlan, P. Verghese, J. Malik, An information maximization model of eye movements, in: L.K. Saul, Y. Weiss, L. Bottou (Eds.), Advances in Neural Information Processing Systems, vol. 17, MIT Press, Cambridge, MA, 2005, pp. 1121–1128.
[49] J. Rissanen, Modeling by the shortest data description, Automatica 14 (1978).
[50] B. Rogoff, J. Mistry, A. Goncu, and C. Mosier, Guided participation in cultural activities by toddlers and caregivers, Monographs of the Society for Research in Child Development, vol. 58 (7, Serial No. 236), 1993.
[51] S. Schaal, A. Ijspeert, A. Billard, Computational approached to motor learning by imitation, Philos. Trans. R. Soc. London B 358 (2003) 537–547.
[52] M. Schlesinger, D. Parisi, The agent-based approach: a new direction for computational models of development, Dev. Rev. 21 (2001) 121–146.
[53] T. R. Shultz, F. Rivest, L. Egri, J. Thivierge. Knowledge-based learning with KBCC, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[54] E. Simoncelli, B. Olshausen, Natural image statistics and neural representation, Ann. Rev. Neurosci. 24 (2001) 1193–1216.
[55] L.B. Smith, Why development matters, in: International Conference on Development and Learning, Bloomington, Indiana, June 2006.
[56] E. Spelke, Initial knowledge: six suggestions, Cognition 50 (1994) 431–445.

[57] N. Sprague, D. Ballard, Eye movements for reward maximization, in: S. Thrun, L. Saul, B. Scholkpf (Eds.), Advances in Neural Information Processing Systems, vol. 16, MIT Press, Cambridge, MA, 2003.
[58] R. Sutton, A. Barto, Reinforcement Learning, MIT Press, Cambridge, MA, 1998.
[59] M. Tomasello, The Cultural Origins of Human Cognition, Harvard, Cambridge, MA, 1999.
[60] J. Triesch, C. Teuscher, G. Deák, E. Carlson, Gaze-following: why (not) learn it?, Dev. Sci. 9 (2006) 125–147.
[61] V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[62] J. Watson, Smiling, cooing, and the game, Merrill-Palmer Quart. 18 (1972) 323–339.
[63] W. Yoshida, S. Ishii, Resolution of uncertainty in prefrontal cortex, Neuron 50 (5) (2006) 781–789.

Gedeon Deak (right) is an Associate Professor of Cognitive Science and Human Development at the University of California at San Diego, where he has been on the faculty since 1999. He received his BA from Vassar College (1990) and his Ph.D. from the Institute of Child Development at the University of Minnesota (1995). He was on the faculty at Vanderbilt University from 1995 to 1999. As director of the Cognitive Development Laboratory (http://www.cogsci.ucsd.edu/~deak/cdlab/), he studies how young children learn and use language and how they solve problems, especially in social contexts like preschools. He also studies infant social-cognitive development using experiments, naturalistic observation, and simulations with computers and robots. This work has been described in over 30 peer-reviewed articles and chapters. He has received research grants from the National Science Foundation, Spencer Foundation, M.I.N.D. Institute at UC-Davis, National Alliance for Autism Research, Kavli Institute for Mind and Brain, and Nicholas Hobbs Foundation. He has received a National Academy of Education Postdoctoral Fellowship and a UC Hellman Junior Faculty Fellowship. He has been a co-organizer of the International Conferences on Development and Learning, 2004–2006. He serves on the editorial board of the *Journal of Cognition and Language* and on the IEEE CIS Technical Committee for Autonomous Mental Development.

Marni Stewart Bartlett is Associate Research Professor at the Institute for Neural Computation, UCSD, where she co-directs the Machine Perception Lab. She studies learning in vision, with application to face recognition and expression analysis. She has authored over 30 articles in scientific journals and refereed conference proceedings, as well as a book, *Face Image Analysis by Unsupervised Learning*, published by Kluwer in 2001. Dr. Bartlett obtained her Bachelor's degree in Mathematics in 1988 from Middlebury College, and her Ph.D. in Cognitive Science and Psychology from University of California, San Diego, in 1998. Her thesis work was conducted with Terry Sejnowski at the Salk Institute. She has also published papers in visual psychophysics with Jeremy Wolfe, neuropsychology with Jordan Grafman, perceptual plasticity with V.S. Ramachandran, machine learning with Javier Movellan, automatic recognition of facial expression with Paul Ekman, cognitive models of face perception with Jim Tanaka, and the visuospatial properties of faces and American Sign Language with Karen Dobkins. She has given numerous invited symposia to international audiences in both machine vision and cognitive neuroscience. Dr. Bartlett has been highly active in organizing workshops and conferences in the areas of learning in vision and affective computing. She co-organized two major conferences, the 11th European Conference on Visual Perception, Bristol, England, and the 3rd International Conference on Development and Learning in San Diego, California, as well as numerous workshops at conferences such as Advances in Neural Information Processing Systems, and International Conference on Computer Vision.

Tony Jebara is an Assistant Professor of Computer Science at Columbia University. He is Director of the Columbia Machine Learning Laboratory whose research focuses upon machine learning, computer vision and related application areas such as human–computer interaction. Jebara is also a Principal Investigator at Columbia's Vision and Graphics Center. He has published over 30 papers in conferences and journals including NIPS, ICML, UAI, COLT, JMLR, CVPR, ICCV, and PAMI. He is the author of the book Machine Learning: Discriminative and Generative (Kluwer). Jebara is the recipient of the Career award from the National Science Foundation and has also received honors for his papers from the International Conference on Machine Learning and from the Pattern Recognition Society. He has served as chair, program committee member, and reviewer for various conferences and workshops. Jebara's research has been featured on television (ABC, BBC, New York One, TechTV, etc.) as well as in the popular press (Wired Online, Scientific American, Newsweek, Science Photo Library, etc.). Jebara obtained his Bachelor's from McGill University (at the McGill Center for Intelligent Machines) in 1996. He obtained his Master's in 1998 and his Ph.D. in 2002 both from the Massachusetts Institute of Technology (at the MIT Media Laboratory). He is currently a member of the IEEE, ACM, and AAAI. Professor Jebara's research and laboratory are supported in part by the Intelligence Community, Microsoft, Alpha Star Corporation, and the National Science Foundation.