The Social Scaffolding of Machine Intelligence

Paul R. Smart

Electronics & Computer Science
University of Southampton
Southampton, UK
Email: ps02v@ecs.soton.ac.uk

Aastha Madaan

Electronics & Computer Science
University of Southampton
Southampton, UK
Email: madaan.aastha@gmail.com

***Abstract*—The Internet provides access to a global space of information assets and computational services. It also, however, serves as a platform for social interaction (e.g., Facebook) and participatory involvement in all manner of online tasks and activities (e.g., Wikipedia). There is a sense, therefore, that the Internet yields an unprecedented form of access to the human social environment: it provides insight into the dynamics of human behavior (both individual and collective), and it additionally provides access to the digital products of human cognitive labor (again, both individual and collective). Such access is interesting from the standpoint of research into machine intelligence, for the human social environment looks to be of crucial importance when it comes to the evolutionary and developmental origins of the human mind. In the present paper, we develop a theoretical account that sees the Internet as providing opportunities for online systems to function as socially-situated agents. The result is a vision of machine intelligence in which advanced forms of cognitive competence are seen to arise from the creation of a new kind of digital socio-ecological niche. The present paper attempts to detail this vision with respect to the notion of socially-scaffolded cognition. It also describes some of the forms of machine learning that may be required to enable online systems to press maximal cognitive benefit from their new-found informational contact with the human social world.**

Keywords*–internet; social intelligence; language; machine intelligence; machine learning.

I. Introduction

There can be little doubt that the Internet represents a milestone in human technical achievement. As a technical accomplishment, the Internet stands testament to our species' capacity for invention, innovation and complex problem-solving. But in this respect, it seems that our species is unique. No other form of terrestrial life is able to build a global communication system, observe the distant reaches of the cosmos, or plumb the murky depths of mathematical mysteria. (Neither, for that matter, are they able to contemplate their species-specific cognitive character and serialize their thoughts in the form of an academic paper!) In a cognitive sense, therefore, we humans are clearly special, for it is only the anatomically modern human mind that has managed to scale the lofty heights of the cognitive mountain. But in being special, we are also very much alone. Cetaceans, chimps and cephalopods are all capable cognizers; but none are in a position to challenge the cognitive supremacy of our own species. The cognitive world, it seems, is bit around a wall that separates two ostensibly distinct cognitive kinds. On one side of that wall we find ourselves; on the other, we find the rest of terrestrial life.

The extent to which we will always be alone (or, indeed, special) is, of course, a moot point. From our vantage point in the cognitive eyrie, we are currently seeking to understand the forces and factors that make our human minds materially possible. And in the light of such understanding, we are striving to build machines in our own cognitive image. It is this undertaking—the traditional focus of Artificial Intelligence (AI) and machine intelligence research—that is perhaps the most difficult of our technical undertakings. Despite some notable successes, the attempt to engineer advanced forms of machine intelligence—machines that emulate our own distinctive forms of cognitive competence—remains, for the most part, a work in progress. The route to human-level intelligence, it seems, is not straightforward. And perhaps this is why we humans find ourselves alone atop the cognitive mountain—the solitary surveyors of the low-lying cognitive terrain.

In the present paper, we wish to consider a particular path up the cognitive mountain. It is a path that focuses attention on the role of the Internet in supporting the emergence of machines with human-like cognitive capabilities. The general aim is perhaps best captured in the form of a question: What impact (if any) does the Internet have on the attempt to engineer machines with human-level intelligence?

There are, to be sure, a number of ways that we might respond to such a question. We might, for example, point to the way in which the Internet has yielded a superabundant supply of widely available digital data, such as image, text and video resources. Such resources have arguably shaped the course of AI research, stimulating research into new forms of machine learning (such as those being explored by Google DeepMind). There is also a sense in which the Internet has played something of an indirect role in advancing the cause of AI. We might, for example, point to the way in which the Internet has yielded a superabundant supply of money for major technology vendors, leading to eye-watering levels of investment in AI-related research.

A different kind of response to the question of how the Internet relates to machine intelligence is to be found in one of our earlier papers [1]. In the context of that paper, we suggested that the Internet, or at least a specific component of the Internet—namely, the Social Web—is poised to yield state-of-the-art advances in machine intelligence. The basis for such optimism was to be found in the (perhaps rather inchoate) claim that the Internet provides a form of informational contact with the *human social environment*, where the notion of the human social environment was cast as the realm in which human behavior (both individual and collective) occurs. As a result of such contact, we suggested that the Internet provides opportunities for machines to observe and interact with humanity, as well as exploit the products of human cognitive and epistemic labor. Thus construed, the Internet was seen to support the emergence of a new kind of cognitively-potent informational ecology: a socio-ecological niche, pregnant with cognitive opportunity.

The present paper introduces a number of extensions to our earlier paper based on the comments and feedback we received from the academic community. These extensions reflect both a broadening and a narrowing of scope. The scope is broadened in the sense that we focus on the Internet rather than just the Social Web. This particular shift in focus is probably of little consequence, for the term "Internet" in the present paper is intended as a catch-all term that encompasses a multiplicity of Internet-related technologies. This includes the World Wide Web, as well as the Web's more specific instantiations, such as the Social Web.

The other shift of focus—the one involving a narrowing of scope—is perhaps more significant. In the present paper, we restrict our attention to a particular kind of cognition, one that goes by the name of *socially-scaffolded cognition*. The meaning of this term will become clearer throughout the course of the paper (see Section II). For present purposes, however, socially-scaffolded cognition can be viewed as a form of cognition whose origins depend on the properties of a social environment (or aspects thereof). This conception is broadly consistent with the flavor of existing work, which links the notion of scaffolded cognition to the acquisition of particular forms of cognitive competence [2]. As we shall see, however, the precise meaning of the term "scaffolded cognition," and its status as an independent cognitive kind, distinct from, say, the likes of extended or embedded cognition, is still a matter that is open to philosophical debate (and disagreement). In this respect, the account of scaffolded cognition offered in Section II represents an attempt to more clearly delineate the notion of scaffolded cognition and distinguish it from other, ostensibly similar, cognitive kinds. This reflects one of the ways the present paper contributes to the philosophical and cognitive scientific literature.

In addition to changes in scope—manifested as the loosening of technological constraints and the tightening of cognitive bonds—the present paper offers a more detailed exposition of some of the mechanisms that may allow certain kinds of intelligent system to press maximal cognitive benefit from Internet-mediated forms of informational contact with the human social world. Relative to the original paper, this particular extension corresponds to neither a constriction nor a dilation of scope. It is, instead, an attempt to highlight the relevance of existing research to socially-scaffolded forms of machine intelligence. Such a move could be seen as marking the first tentative steps towards a concrete empirical research agenda, with perhaps potentially profound implications for the development of (e.g.) cognitive computing systems. For the most part, however, our aim in the present paper is to establish the theoretical basis for research in this area. We thus focus our attention on the following issues:

1) Why is the notion of socially-scaffolded cognition relevant to AI research?
2) How is the Internet apt to function as a cognitively-potent informational ecology—one that supports the emergence of advanced forms of machine intelligence?

In addressing these issues, we attempt to draw on ideas, insights and empirical findings that are strewn across a rich array of academic disciplines. This is, of course, a high-risk strategy: In taking the transdisciplinary path, one often embarks on a treacherous journey into the intellectual wilderness—an interdisciplinary no man's land where the intellectual payoff is, at best, uncertain and the reputational rewards (relative to one's academic career) are probably zero. So be it. Inasmuch as the following is deemed to yield little in the way of a genuine advance in our understanding of how to build an intelligent machine, we will at least take comfort in the fact that we have saved someone else the journey.

The structure of the paper is as follows: Section II focuses on the notion of scaffolded cognition and develops a developmentally-oriented conceptual account that distinguishes scaffolded cognition from ostensibly similar cognitive kinds, such as extended cognition and embedded cognition. Section III seeks to highlight the relevance of social forces and factors to human intelligence. It does this by describing the way in which the human mind is shaped by the human social environment, in both an evolutionary and an ontogenetic sense. Such insights provide the basis for Section IV, which discusses the way in which the Internet allows certain kinds of AI systems—dubbed *social machines*—to function as socially-situated or socially-embedded agents. Section V surveys a number of different forms of machine learning, namely, social, active, language, predictive and incremental learning. The aim here is to identify some of the cognitive prerequisites for a social machine—the capabilities that enable a social machine to press cognitive benefit from its informational contact with the human social environment. Finally, in Section VI, we summarize some of the core ideas discussed throughout the paper and mention some areas for future theoretical and empirical work.

II. Scaffolded Cognition

One way of understanding the significance of the human social environment to the development of human cognitive capabilities (in both a phylogenetic and an ontogenetic sense) is via the notion of *scaffolded cognition* [2][3]. The term "scaffolded cognition" is typically (although not always) used to refer to a cognitive ability that emerges as the result of an agent's exposure to scaffolding resources, where the resources in question form part of the environment of a cognitive agent and play an active role in shaping the agent's cognitive development. In the case of socially-scaffolded cognition, such resources are most obviously thought of as other human individuals, as well as perhaps the products of human cultural innovation (i.e., artifacts, knowledge, norms, language, tools, practices, and so on). It is these resources—the material elements of what we call the human social environment—that help to shape the course of human cognitive development and the trajectory of human cognitive evolution. In an important sense, it is our exposure (and response) to such resources that makes us what we are—a species able to negotiate cognitive terrains that lie beyond the ken of other earthly critters.

Our aim in the present paper is to apply the notion of socially-scaffolded cognition to the realm of AI systems. In particular, we suggest that the path to state-of-the-art advances in machine intelligence may be revealed by a consideration of the ways in which various forms of social scaffolding shape the course of human cognitive development and (over longer timescales) the course of human cognitive evolution. Before we begin to unpack this claim, however, it will help to have a clearer understanding of what is meant by the notion of scaffolded cognition. This is important, because the term "scaffolded cognition" is one that is used in different ways within the cognitive scientific literature. The term is perhaps most often encountered in the context of educational or developmental psychology, but this is not always the case. In addition, scaffolded cognition is only one among a number of cognitively-oriented concepts that have been the focus of cognitive scientific attention, and the relationships between these concepts are, it has to said, not fully understood. It is easy, for example, to think of scaffolded cognition as denoting a particular kind of cognition, i.e., a *cognitive kind*. But much the same could be said about other elements of the cognitive scientific lexicon, including the notions of extended [4]–[6], embedded [7], situated [8], distributed [9] and embodied cognition [10]. Distinguishing between these concepts and identifying their relationships to one another is a major theoretical undertaking, and it is not one that we can hope to achieve in the present paper. That said, there is a particular need to understand the distinction between scaffolded cognition and at least some other cognitive kinds, most notably extended cognition and embedded cognition. This is important, because the relevance of the Internet to issues of machine intelligence has been the focus of previous work. In particular, Smart [11] has suggested that the Internet provides opportunities for both extended and embedded forms of machine intelligence, and the human social environment is deemed to be relevant in both cases. At a minimum, therefore, we need to understand how (and to what degree) the appeal to scaffolded cognition provides us with a novel view of the Internet and its contributions to machine intelligence.

Perhaps the most significant obstacle to a successful discrimination between scaffolded cognition and other cognitive kinds lies in the fact that appeals to scaffolded cognition are sometimes encountered outside of a developmental context. Clark [12], for example, discusses the way in which a variety of external scaffolds, including public language and culture, help to "mold and orchestrate behavior" in adaptive or strategic ways [12, pp. 32–33]. Relative to Clark's vision of the human mind as an extended cognitive organization—one in which cognitive processing routines rely on a multiplicity of resources drawn from the biological, social and technological realms—it is perhaps easy to see scaffolded cognition as related to the notion of extended cognition. This is especially so if scaffolding resources are deemed to play a role that goes beyond the mere causal dependence of cognitive processing routines on aspects of the external environment. In fact, something along these lines is suggested by Arnau et al. [13], as part of their attempt to define scaffolded cognition:

> Scaffolded cognition is the idea that (at least some of) our cognitive capacities both depend on and have been transformed by our manipulation of environmental resources. The claim here is not about mere causal dependence, but about integrative coupling between internal and external elements. [13, p. 56]

What is crucial here is the claim that scaffolded cognition involves something more than "mere causal dependence." One way of making sense of this claim is via an appeal to the distinction between causal and constitutive relevance [14][15], where the notion of constitutive relevance implies that some resource is an intrinsic element of the physical fabric that realizes or constitutes some phenomenon of interest. (This contrasts with the more familiar notion of causal relevance, where some resource is seen to cause the occurrence of some phenomenon.) In the case of extended cognition, the claim is that some extra-organismic resource is of constitutive relevance to some cognitive phenomenon (e.g., a given cognitive process), and it therefore forms part of the physical machinery that realizes the cognitive phenomenon in question. One way of thinking about scaffolded cognition, therefore, is to see it as a form of cognition in which some form of extra-organismic scaffolding resource is constitutively relevant to cognition.

The problem with this proposal is immediately obvious: By appealing to issues of constitutive relevance, the distinction between extended and scaffolded cognition is obscured, perhaps to the point where the two concepts are indistinguishable. In this sense, the appeal to "integrative coupling" in the aforementioned quote by Arnau et al. is problematic because it resembles similar appeals to integrative coupling made in respect of extended cognition (see [16]).

In the interests of distinguishing scaffolded cognition from extended cognition, we may opt to drop our allegiance to constitutive relevance and recast the relevance relation as one of causal relevance. In other words, when we reflect on the relation between some scaffolding resource and a particular cognitive capacity, it may make sense to view the relation in purely causal terms. The problem, in this case, is that by embracing the notion of causal relevance we are in danger of confusing scaffolded cognition with another kind of cognition, namely, embedded cognition [7]. Like extended cognition, embedded cognition recognizes the dependence of cognitive properties on elements that lie external to the cognitive system. In this case, however, the dependence relation is best understood with respect to the notion of causal relevance rather than constitutive relevance.

The upshot of all this is a dilemma that turns on the relationship between the properties of a cognitive system and the role played by some extra-systemic resource. If we conceive of this relationship in such a way that the resource is constitutively relevant to the properties of a cognitive system, then scaffolded cognition emerges as nothing more than extended cognition. On the other hand, if we drop the appeal to constitutive relevance and instead conceive of the relationship from the standpoint of causal relevance, then what we are left with is nothing other than embedded cognition. Either way, the notion of scaffolded cognition looks to be conceptually redundant.

Our approach to the resolution of this dilemma is rooted in an appeal to the notion of *mechanistic explanation*, which is a form of explanation that focuses on the mechanisms that are deemed to be responsible for some phenomenon of interest (e.g., [17]). Two kinds of mechanistic explanation look to be of particular importance when it comes to extended and embedded cognition. These are *causal mechanistic explanation*, which is tied to claims of causal relevance, and *constitutive mechanistic explanation*, which is tied to claims of constitutive relevance (see [15]). One way to think about the distinction between embedded and extended cognition is thus to see embedded and extended cognitive systems as the targets of different kinds of mechanistically-oriented explanatory account: causal mechanistic explanations are thus best suited for embedded cognition, while constitutive mechanistic explanations are best suited for extended cognition (see [18]).

So much for embedded and extended cognition. But what about scaffolded cognition? Is there a particular kind of mechanistic explanation that is apt for scaffolded cognition, and is this form of explanation sufficiently distinct from mechanistic explanations of the causal or constitutive stripe?

In fact, we suggest that scaffolded cognition is best approached from the perspective of what are called *developmental explanations* [15][19]. Such explanations seek to trace the historical lineage of a phenomenon, helping us understand how the phenomenon relates to an interacting nexus of causally-active historical forces and factors. Developmental explanations are, in essence, an attempt to detail the causal history of phenomena, and, in this sense, they bear much in common with the sorts of explanations encountered in the historical sciences [20]. (Indeed, we regard historical explanation as a particular form of developmental explanation.) As with causal and constitutive mechanistic explanations, developmental explanations are typically cast as a particular form of mechanistic explanation. In other words, developmental explanations resemble other forms of mechanistic explanation in that their explanatory heft inheres in the attempt to provide a complete description of the mechanisms that are responsible for some target phenomenon. It is perhaps tempting to refer to such mechanisms as *developmental mechanisms*, although the previous use of this term is mostly confined to the realms of developmental biology [21]. For present purposes, we will use the terms "developmental explanation" and "developmental mechanism" in a generic sense to refer to explanations and mechanisms that are encountered in a variety of disciplinary contexts. These include developmental biology (ontogenetic explanations), history (historical explanations), and evolutionary biology (evolutionary explanations).

We have already seen how issues of constitutive/causal explanation and constitutive/causal relevance can be used to discriminate between extended and embedded cognition. The same approach, we suggest, can be used to inform our theoretical understanding of scaffolded cognition. Scaffolded cognition can thus be thought of as a particular form of cognition (i.e., a cognitive kind) that is the apt target of a particular kind of explanatory approach. Just as constitutive explanations are appropriate to extended cognition, and causal explanations are appropriate to embedded cognition, so developmental explanations, we suggest, are appropriate to scaffolded cognition. As with extended and embedded cognition, scaffolded cognition involves resources that are located external to the boundaries of a cognitive system. In the case of scaffolded cognition, however, these resources are deemed to be of *developmental relevance* with regard to whatever cognitive phenomenon is the target (i.e., the explananda) of mechanistic explanation. In fact, these resources are nothing other than what we have been calling scaffolding resources. In essence, scaffolding resources are the material elements of developmental mechanisms that are described by developmental explanations. Thus construed, developmental relevance can be regarded as a particular kind of explanatory relevance. Just like causal and constitutive relevance (which are also forms of explanatory relevance), developmental relevance highlights the relevance of some extra-systemic resource for the purposes of (developmentally-oriented) mechanistic explanations.

There is, of course, a sense in which developmental explanations are similar to causal explanations, and this might be seen to serve as a source of confusion when it comes to the distinction between scaffolded cognition and embedded cognition. Just like causal explanations, developmental explanations reveal the ways in which a set of causally-active antecedent forces and factors conspire to yield some sort of outcome. There is, however, no reason why this should lead to confusion between the notions of scaffolded and embedded cognition. Embedded cognition seeks to explain cognitive phenomena with respect to causal influences that operate in the here-and-now, and, as a result, the resources picked out by the notion of causal relevance are always ones that are present in the local environment of a cognitive system. Such need not be the case with scaffolded cognition. In the case of scaffolded cognition, the relevant resources (scaffolding resources) need not be present in the local environment of a cognitive system. Indeed, in some cases, such resources may no longer exist. Issues of spatial and temporal proximity are thus of crucial importance for embedded cognition, but they are of relatively little importance for scaffolded cognition.

This is also, as it happens, the reason why developmental explanation/relevance cannot be equated with constitutive explanation/relevance. The resources that are relevant to constitutive explanations are always deemed to be physically present because they form part of the mechanism that realizes occurrent cognitive phenomena. Such is not the case with scaffolding resources, and it is for this reason that the concept of scaffolded cognition cannot be equated with extended cognition.

Having said all this, it should be noted that the relationship between the notions of causal, constitutive and developmental relevance is still something that is up for grabs. Ylikoski [15], for example, suggests that developmental explanations are at times complex amalgams of constitutively and causally relevant factors. Developmental explanations, he suggests, sometimes involve complex forms of reciprocal causation and mutual influence that are difficult to track with simple causal accounts. The result, it seems, is that developmental explanations are not purely constitutive explanations, but neither are they purely causal explanations. Perhaps it is this inextricable entanglement between constitutive and causal relevance, spread out over (sometimes significant) periods of time, that best accommodates the intuition that scaffolded cognition relies on something more than "mere causal dependence" and involves some degree of "integrative coupling" [13].

The result of all this is a conception of scaffolded cognition that appeals to the notions of developmental explanation and developmental relevance. Scaffolded cognition is, in essence, a developmentally-oriented concept. One virtue of this developmentally-oriented conception is that it is nicely aligned with the bulk of research into scaffolded cognition, most of which has been undertaken in an educational or developmental context. The conception also, however, provides a means of discriminating scaffolded cognition from extended and embedded cognition, and it does so in such a way as to (perhaps) reveal why these concepts have proved so hard to disentangle.

III. The Social Origins of Human Intelligence

The form of scaffolded cognition that concerns in the present paper is referred to as *socially-scaffolded cognition*.

This is a form of scaffolded cognition that is distinguished with respect to the nature of the resource that does the scaffolding (see [2]). In the case of socially-scaffolded cognition, we are interested in situations where the social environment (or some aspect thereof) plays a role in the developmental emergence of cognitive capabilities.

Why should socially-scaffolded cognition be of any interest or relevance to those interested in machine intelligence? To answer this question, it will help to introduce two substantive strands of empirical and theoretical research: one focused on the evolution of the human cognitive system; the other, on the ontogenetic development of human cognitive capabilities.

Let us first direct our attention to issues of human cognitive evolution. There is clearly something special about human cognition—something that makes our own species cognitively unique. But what is it about the evolutionary history of our species that accounts for this remarkable divergence in cognitive power and sophistication?

One response to this question focuses on the selection pressures arising from the physical environment—the need to find food, ward off predators, deal with climatic changes, and so on. The inadequacy of this account is immediately obvious: it fails to explain what it is that gives human cognition its rather distinctive flavor. Why is it that humans—and only humans—are in possession of advanced cognitive capabilities? Presumably, we humans are not alone in having to deal with a range of ecological challenges, so why has evolution not driven other forms of terrestrial life to evolve a similar cognitive profile?

A different (albeit related) response focuses on the challenges thrown up by a particular kind of ecological niche: the human (or, perhaps better, hominin) social environment. According to approaches of this ilk, the forces and factors that account for the evolutionary emergence of the modern human mind are not to be found in the physical environment of our hominin ancestors. Instead, it is suggested that the well-spring of human cognitive success is to be found in the socio-ecological realm. The idea, in essence, is that the human mind evolved to deal with the vagaries of an environment that was itself constituted by other humans (and their minds). In other words, we humans 'created' the specific socio-ecological niche from which the modern human mind emerged. The human mind, according to this view, corresponds to something of a socially-created artifact—a device that was forged in a crucible of our own creation.

This idea actually comes in a variety of flavors. It surfaces, in somewhat different forms, in a number of recent evolutionary hypotheses, including the social brain hypothesis [22], the Machiavellian intelligence hypothesis [23], the cultural intelligence hypothesis [24], the social intelligence hypothesis [25], the sexual selection hypothesis [26], and the ecological dominance-social competition hypothesis [27]. What unites these hypotheses is a claim about the evolutionary significance of intra-specific social competition. Flinn [28], for example, suggests that:

> ...the human psyche was designed primarily to contend with social relationships, whereas the physical environment was relatively less important. Most natural selection in regard to brain evolution was a consequence of interactions with conspecifics, not with food and climate...To a degree that far surpasses that of any other species, human mental processes must contend with a constantly changing information environment of their own creation. (pp.73–74)

There are two aspects to this socially-oriented account of human cognitive evolution that are worth highlighting. The first is the emergence of a form of runaway directional selection in which the emergence of cognitively, socially, and behaviorally sophisticated individuals merely serves to exacerbate existing social selection pressures, increasing the complexity of the social environment that evolution must contend with. We thus encounter something of a positive feedback loop, in which individual cognitive sophistication leads to greater social complexity, which in turn leads to ever-greater demands for cognitive sophistication. The result is something of a red queen situation (see [29]): cognitive sophistication begets social complexity, which in turn intensifies the drive toward cognitive sophistication. For the sake of convenience, let us dub this *the red queen of socio-ecological complexity*.

The mechanism that lies at the heart of this particular red queen is well-documented. Its lineage can be traced to Humphrey [30] who cast the feedback loop as a form of ratchet, a "self-winding watch to increase the general intellectual standing of the species" (p. 311). This is a useful metaphor, in the sense that it helps us see human cognitive evolution as something of an autocatalytic process—one in which a particular form of cognitive-evolutionary progress lays the foundation for yet further increments in cognitive power and sophistication [27]. The metaphor is also useful in that it draws attention to the dynamic nature of the human social environment. In contrast to the physical environment, whose features are, for the most part, relatively enduring (or at least predictable from one generation to the next), the topography of the social terrain is riven by the tectonic forces of cognitive-evolutionary change.

There is, however, another aspect to this theoretical account that highlights the rather unique nature of the human social environment. In addition to being a dynamically changing environment—one whose complexity tracks progressive increases in human cognitive sophistication—the human social environment is also one that is shaped by the forces of cultural evolution. This obviously adds to the unpredictability of social environments across inter-generational timeframes; but it also (and perhaps more importantly) lays the foundation for the diversification of social environments *within* any given generation. The upshot is that, from the standpoint of evolution, the human socio-ecological niche is one whose features are, at best, difficult to predict. Faced with such a situation, there is perhaps little that evolution can do except yield organisms that are equipped with a combination of powerful learning mechanisms and extraordinary levels of (cognitive) phenotypic plasticity:

> Once human cultural evolution began to accelerate and languages began to change rapidly, there would have been strong selection for general and language-specific increases in brain plasticity. Since the one thing that is consistently stable in a rapidly changing culture is the culture's context-dependent flexibility (which the cultural evolutionary process itself creates), there is persistent selection for increasingly flexible and sophisticated ways of learning, including language-learning. [31, p. 2154]

It is this particular form of phenotypic plasticity that is the source of a second red queen, one that we will dub *the red queen of socio-ecological variability*. To help us understand this particular red queen, note that phenotypic plasticity, when situated within the specific context of cultural innovation and learning, is the source of a second ratchet-like mechanism that drives the evolution of ever-greater levels of plasticity. The reason for this is that phenotypic plasticity goes hand-in-hand with phenotypic variability, and phenotypic variability contributes to precisely the sorts of socio-cultural differentiation that make the human socio-ecological niche so hard for evolutionary processes to predict. The result is a positive feedback loop, in which socio-ecological variability drives the evolution of cognitive plasticity, which in turn gives rise to ever-more opportunities for socio-ecological diversification.

It is at this point that our attention begins to switch from phylogeny to ontogeny. For the complex and capricious nature of the human social environment may be useful in helping us understand some of the characteristic features of human ontogeny. The first of these concerns the plasticity and structural lability of the human brain, which looks to be particularly pronounced in human infants and young adults (e.g., [32]). The second is the protracted nature of human maturational processes (i.e., the period of extended development known as childhood). Both these features can be regarded as adaptations to a complex and inconstant human social environment. The developmental profile of the human brain, for example, may provide the basis for extreme forms of neural plasticity, in which the structural and functional architecture of the biological brain adapts to the demands thrown up by a complex and unpredictable environment [33][34]. Similarly, the functional significance of extended development (or childhood) is typically cashed out in terms of the opportunities for learning. In particular, it has been suggested that an extended period of development is required to enable human individuals to learn the skills required in later life [35]. Interestingly, human infants are born helpless, and they remain immature for longer than might be expected [36]. It is easy to regard this period of neonatal altriciality as something of a costly encumbrance that is imposed on infants (and parents), and which hampers the opportunities for subsequent social learning. Note, however, that in being helpless, the human infant is totally dependent on her human care-givers, and this establishes the basis for particularly intimate forms of social interaction—the very stuff that drives socially-scaffolded development. In this sense, the altricial status of young human infants, while easily glossed as something maladaptive and costly, can also be seen to lay the foundation for future forms of socially-directed or socially-inflected learning. As noted by Nelson [37], this period of "enforced dependent sociality is both the foundation for the social mind of humans and for the particular course of social-cognitive development found in the human child" (p. 367).

The social environment plays an important role in shaping human cognitive development throughout childhood, and often well into adulthood. And it is here that we find the bulk of research into scaffolded cognition. One of the foremost proponents of socially-scaffolded development is the Soviet psychologist, Lev Vygotsky. Vygotsky argued that the nature of our interaction with socially-significant others holds the key to understanding human cognition (see [38]). Human intelligence is, according to Vygotsky, something that emerges as a result of social interactions with other human beings.

The upshot of all this is a vision of the human mind that is thoroughly social, both in origin and in orientation. We have seen how the ever-changing topography of the social terrain plays a crucial role in shaping the trajectory of human cognitive evolution, and we have also seen how social interactions are poised to play a crucial role in cognitive development. Being social, it seems, is what *makes* us human. Irrespective of whether our attention is focused on issues of phylogeny or ontogeny, the human social environment emerges as of crucial importance in our attempts to understand the developmental mechanisms that give rise to that most marvelous, and yet most mysterious, of cognitive devices: the modern human mind.

IV. Social Machines

Inasmuch as we see the human mind as a socially-created artifact—or a socially-engineered cognitive machine—then perhaps a consideration of social forces and factors is relevant to our ongoing effort to develop AI systems. Perhaps, in other words, a consideration of the social realm enables us to trace a path to the top of the cognitive mountain—a path that was followed (and in some sense forged) by our own species. It is, no doubt, a precarious and ill-defined path, one whose course is punctuated by soaring cliffs and gaping chasms. But it is, nevertheless, a path. And, given our solitary status at the top of the cognitive hierarchy, it may very well turn out to be the only path available.

There is, of course, nothing new in the idea that a consideration of the social environment is relevant to the effort to build intelligent machines. The idea is, in fact, the mainstay of the field of social and (to a lesser extent) developmental robotics [38]–[41], both of which emphasize the role of social interaction and engagement in the development of advanced cognitive capabilities. Consider, for example, the following quotes from Kerstin Dautenhahn and colleagues, both of which appeal to ideas rehearsed in the previous section (see Section III):

> If social intelligence, in evolutionary terms, 'came first' in the development of primate intelligence, and then later was applied to other domains, then one may extrapolate and apply this 'evolutionary history' to machines, too. Accordingly, from an evolutionary perspective, then intelligent robots need to be socially intelligent robots. [41, p. 295]

> Our research is based on the assumption that in order to study the cognitive development of robots we have to consider the 'robot in society', i.e., using Vygotsky's approach to see social interactions as fundamental, and as a context which can scaffold the development of cognitively richer functionalities. [42, p. 6]

The theoretical position proposed in the present paper is based on precisely these sorts of assumptions. The only significant difference is the nature of the system that is seen to be the beneficiary of socially-scaffolded development. In the case of social robotics, of course, the relevant system is typically some form of robot, typically one that is implemented as a physical entity, equipped with a real 'body' that serves as the basis for robot–human and sometimes robot–robot interactions. The systems of interest in the present paper are somewhat different. We are interested in a class of systems that operate in the online realm of the Internet and which typically exist in the form of computer programs. In this sense, the kinds of intelligent system that we are interested in would no doubt be regarded as 'disembodied' and thus incapable of functioning as socially-situated agents. While we do not wish to contest the claim that some sort of distinction should be made between a purely online system and a real-world robot, it is not clear that the notion of embodiment is best placed to motivate this distinction. A fuller discussion of this issue would take us too far afield; however, it is worth noting that some kinds of online system might be regarded as embodied by virtue of the forms of real-world sensory/motor contact provided by (e.g.) the Internet of Things (IoT) [11]. It is also worth noting that some have questioned the extent to which issues of embodiment apply *only* to the realm of real-world, physical systems, as opposed to their purely virtual counterparts [43][44].

In any case, our primary interest in the present paper is the extent to which the Internet enables AI systems to be embedded or situated within the human social environment. Central to this idea is the claim that the Internet provides a form of informational contact with the human social environment. This particular claim will probably require little in the way of a detailed defense, for the Internet has undoubtedly provided a rich array of opportunities for conventional computational systems to engage with human agents and observe their behavior at both an individual and collective level. The Social Web is just one example of this. With the advent of social networking sites, microblogging services, and media sharing systems, the online environment affords ever-deeper insights into the dynamics of human social behavior [45]. Additional forms of contact are arguably provided by an ever-expanding array of mobile and portable computing devices, Internet-enabled sensors, and IoT devices.

It is, of course, easy to think that this notion of the Internet providing contact with (or access to) the human social environment should be interpreted solely in observational terms, i.e., as the Internet enabling machines to monitor human behavior at both the individual and collective levels. In fact, a somewhat broader notion is in play here. We see the Internet as providing access to an online ecology of human-generated digital assets, some of which indirectly contribute to the (social) shaping of machine-based capabilities. Consider, for example, the way in which the addition of descriptive tags and annotations to a set of image resources assists with the development of machine vision systems [46]. Here, human contributions yield a body of training data that is apt to support a particular kind of machine learning. Such possibilities are explicitly recognized by those who seek to engage human subjects in computationally-difficult tasks. With respect to citizen science systems, for example, Lintott and Reed [47] note that one of the limiting factors in the development of automated processing solutions is the availability of sufficiently well-structured training data sets, and that one of the key advantages of citizen science projects is the provision of such data sets. Similarly, when it comes to a class of systems known as Games With A Purpose (GWAPs), von Ahn and Dabbish [48] are keen to stress the role of human contributions in giving rise to ever-more intelligent (and human-like) forms of machine-based processing:

> By leveraging the human time spent playing games online, GWAP game developers are able to capture large sets of training data that express uniquely human perceptual capabilities. This data can contribute to the goal of developing computer programs and automated systems with advanced perceptual or intelligence skills. [48, p. 67]

The key point, here, is that by virtue of human contributions, a set of digital resources that were previously too ill-structured to support machine learning are transformed into something that is much better aligned with the requirements of machine learning algorithms. Something along these lines also applies to systems such as IBM Watson [49], which benefit from the online availability of socially-generated and socially-structured resources (e.g., Wikipedia). In these cases, advances in machine intelligence derive from the access the Internet provides to the human social environment, but it is not a form of access that can be characterized solely in observational terms.

The basic vision, then, is one of the Internet providing a form of informational access to the human social environment. Relative to this vision, we suggest that the Internet provides opportunities for AI systems to be embedded or situated within the human social environment, enabling them to benefit from various forms of socially-scaffolded development. For the purposes of this paper, we will refer to these socially-situated systems as *social machines*. A social machine is thus a particular kind of intelligent system that benefits (in a cognitive sense) from Internet-mediated forms of informational contact with the human social environment. It is, in essence, a machine whose cognitive capabilities are tied to its status as a socially-situated agent.

V. Socially-Scaffolded Cognition and Machine Learning

Clearly, not every kind of intelligent system is likely to qualify as a social machine. The status of social machines as socially-situated or socially-embedded (see [38][50]) systems perhaps goes some way to limning the relevant class of systems. But even the notion of social situatedness seems somewhat insufficient. Mere exposure to the human social environment will not cause a socially-oriented cognitive critter to develop human-level cognitive abilities. If it did, then we would recognize dogs and budgies as kindred cognitive spirits.

As it stands, therefore, the notion of a social machine remains somewhat vague. It is, in particular, unclear what kinds of intelligent system should be counted as social machines. What are the peculiar features of a social machine that enable it to function as a socially-situated agent? What are the details of its cognitive architecture? What are the ways in which a social machine is poised to benefit from socially-scaffolded development? What is it that enables a social machine to press maximal cognitive benefit from its immersion in a socio-ecological niche? And why should the human social environment (as opposed to any other kind of environment) be of particular relevance to the emergence of advanced cognitive capabilities?

In the present section, we attempt to provide some initial answers to these questions. For the most part, we restrict our attention to the realm of learning. Obviously, there is more to being a social machine than just learning. It may be, for example, that only certain kinds of computational organization are able to fully benefit from the forms of learning detailed below. It is, in addition, at least plausible that only certain kinds of system (e.g., neural networks) are able to exhibit the kinds of fluid, context-sensitive response that are typically associated with intelligent behavior—something that is nicely captured by Clark's [5, p. 107] notion of *intrinsic suitability*. Perhaps, to extend Clark's claims about intrinsic suitability, there is only one kind of computational substructure (a connectionist-style deep learning system, perhaps) that is able to grapple with the complexity of the human social environment and yield something vaguely reminiscent of human-level intelligence. In the interests of space (and, to be honest, the limits of our own intellectual horizons), we will avoid a detailed discussion of these sorts of architectural issues (although see Section V-D for some initial thoughts in this area).

A selective focus on learning seems particularly apt given the developmentally-oriented conception of scaffolded cognition that was proposed in Section II. It is also one that accommodates the discussion in Section III concerning the role of the human social environment in scaffolding the ontogenetic and phylogenetic emergence of the human cognitive system. There is clearly a sense in which learning is perhaps somewhat better suited to accommodate ontogenetic forms of social scaffolding—the forms of scaffolding that occur during the lifetime of a single individual. But perhaps the appeal to learning can also be extended to the domain of evolutionary mechanisms. As is noted by Chalup [51], "[d]uring the time phase of evolution the structure of the genome undergoes a process of phylogenetic learning which is based on evolutionary concepts such as selection and mutation" (p. 448).

In what follows we direct our attention to the following forms of learning: social learning, active learning, language learning, predictive learning and incremental learning. The discussion of these forms of learning reveals what we take to be some of the essential features of a social machine. These include:

- **Phenotypic Plasticity:** Plasticity is clearly the *sine qua non* of a learning system. As such, this feature is relevant to all forms of learning. Issues of plasticity are particularly relevant when it comes to changes in the computational organization of a system as a result of maturational processes. These issues are discussed in the context of incremental learning (see Section V-E).
- **Active Engagement:** A recent focus of cognitive science research is the way in which learners self-structure their learning experiences and thus influence their own learning experiences. These issues are tackled in the context of research into active learning (see Section V-B).

The ensuing discussion is also intended to highlight some of the features of the human social environment that make it of particular interest as both the target of learning and as a context in which learning occurs. These features are perhaps most clearly resolved with regard to the following forms of learning:

- **Social Learning:** The human social environment serves as a source of knowledge that, at least in some cases, can be used to bypass other forms of learning. Such a vision bears a close resemblance to the apprenticeship model of human cognitive evolution, as discussed by Sterelny [52].
- **Predictive Learning:** One of the features of the human social environment is its complexity, which is determined, at least in part, by the cognitive sophistication of its human inhabitants. In negotiating the social environment, a machine learning system must learn to navigate a terrain whose topography is both complex and unstable. The attempt to gain a predictive toehold in this terrain may lead to the emergence of a representational and computational economy that profoundly alters the cognitive repertoire of a social machine.
- **Language Learning:** Finally, we suggest that in dealing with the human social environment, social machines are gifted a particularly potent cognitive tool in the form of language. Such a tool can be seen to magnify other forms of scaffolded development, open up new learning opportunities, and, perhaps most importantly, lay the foundation for profound shifts in cognitive functionality.

A. Social Learning

By virtue of the access it affords to the human social environment, the Internet provides a number of opportunities to observe and monitor different aspects of human behavior. This is important, for we humans are the locus of particular kinds of skill and expertise that reflect our experience with particular domains. Such forms of skill and expertise are typically driven by bodies of knowledge that we have acquired through extensive training and experience, much of which is itself scaffolded by the human social environment. This presents a challenge for the machine-based emulation of human cognitive competence. If human competence develops as a result of the scaffolding provided by a surrounding nexus of social and cultural resources—if, in other words, human capabilities are the products of socially-scaffolded learning experiences—then perhaps it should come as no surprise that human cognitive tasks pose something of a challenge for machine-based systems.

It is here, we suggest, that the Internet provides us with an opportunity to extend the reach of machine-based capabilities. The basic idea is that the Internet can be used as a form of *social observatory*—one that enables machines to observe the human social environment and acquire information about various forms of human competence. From this perspective, the Internet can be seen to support a particular form of *social learning*: it enables us to treat the human social environment as a source of information and knowledge that can be mined and monitored as a means of extending the cognitive and epistemic reach of machine-based systems.

All of this no doubt sounds uncomfortably vague, so let us consider a specific example—one that is admittedly hypothetical yet not so remote as to lie beyond our current technological horizons. The example concerns the effort to develop Autonomous Road Vehicles (ARVs), such as self-driving trucks and cars. ARVs obviously have a range of capabilities, and not all these capabilities are ones that need rely on social learning (at least of the sort we are discussing here). When it comes to an ability to respond to a multitude of driving-relevant situations, however, it looks likely that ARVs will need to possess some of the 'commonsense' knowledge that human drivers have acquired as a result of their experience behind the wheel. Such experience underlies our ability to anticipate the likely behavior of other road users, our ability to behave appropriately at an intersection, our ability to adjust our driving behavior given specific meteorological conditions, and so on. An experienced human driver thus embodies a wealth of knowledge and experience, at least some of which looks to be relevant to the design of autonomous vehicles.

How do we go about building vehicles that possess the behavioral competence and road-related *savoir faire* exhibited by the typical human driver? One option is to enlist the use of conventional knowledge elicitation techniques [53] in order to create formal models of the relevant body of human knowledge. The problem with this approach is that it is likely to require substantial time and effort, especially when one considers the complexity of the target domain, not to mention the diversity of driving practices exhibited by both individuals and cultural groups.

Here is another approach: track the behavior of human-driven vehicles as they move around the road network and then attempt to extract and formalize interesting regularities from the resultant body of 'experiential data'. Such data sets are likely to be particularly valuable in cases where it is possible to track the precise behavior of vehicles at particular locations, such as at an intersection, a roundabout, or a notorious black spot. Additional value comes from the ability to track other kinds of information, such as the use of driver signaling mechanisms (e.g., the use of indicators and headlights) and information about prevailing meteorological conditions (e.g., the presence of fog).

It might, of course, be suggested that human drivers are not the most suitable role models for ARVs, especially if one reflects on the popularity of *The Fast and the Furious* movie franchise. However, even if human behavior should fail to provide a suitable template for machine behavior, it may still be important to learn about such behavior as the basis for anticipating (and responding to) certain situations. This seems particularly relevant to the ARV case, where, in all likelihood, we will encounter a transitional era in which autonomous vehicles are required to share the road with human drivers. There is, in addition, no reason why we should regard the end-product of social learning as being solely about the acquisition of some form of purely behavioral competence. Social learning may thus support the acquisition of knowledge about the unwritten rules of social conduct—the culturally-specific norms, conventions, and practices that shape the dynamics of human social behavior. Social learning thus provides us with a socially- and empirically-grounded approach to what is commonly referred to as machine ethics (also known as machine morality, artificial morality, or computational ethics) [54]. In essence, the idea is to rely on the human social environment to provide insight about the unwritten 'rules' that govern behavior in different social situations. Such knowledge seems particularly important in situations where machines are required to participate in social processes or interact with human agents. Autonomous vehicles are, of course, a case in point. When it comes to the effort to develop ARVs, therefore, social learning may provide a solution to Walport's [55] worry about the need to codify "tacitly accepted 'rules of the road' norms" (p. 24).

The main point of the ARV example is that it helps us see how a particular form of (observational) access to the human social environment can provide insight into bodies of experientially-grounded knowledge, some of which may be relevant to the attempt to engineer a particular kind of intelligent system. The vision is thus one in which advanced forms of machine intelligence come about as the result of a deliberate attempt to learn from the human social environment. According to this vision, machine intelligence is, in a sense, parasitic on human experience: it relies on the experience that humans have in order to short-circuit the acquisition of particular forms of cognitive and behavioral competence, many of which may be hard to acquire via other means.

There is, of course, no reason to think that social learning is restricted to the realm of ARVs. With the advent of the IoT, an increasing number of everyday objects are poised to shed light on the nature of our embodied interactions with a plethora of everyday artifacts, perhaps providing insight into the structure of epistemic actions (see [56]) and culturally-nuanced forms of cognitive practice (see [57]). Needless to say, when it comes to considering the significance of such devices, the emphasis is typically on the way in which issues of network-enablement help or hinder *human* action. But in light of the present discussion, we can perhaps begin to ask ourselves whether there is any reason why such devices should not be used in roughly the same manner as a network-enabled automobile, i.e., as a source of information about the kinds of skill and expertise that might be required to exhibit competence in some otherwise intractable task domain. This is surely a laudable target for machine intelligence research, irrespective of whatever technical challenges confront the effort; for why assume that *a priori* methods can always yield the level of behavioral and cognitive complexity required to deal with domains where human forms of competence only seem to emerge as the result of extensive training and experience?

Based on the foregoing examples, it might be thought that the notion of social learning only applies to situations involving the real-time monitoring of human behavior, as provided, perhaps, by IoT devices. Real-time monitoring, however, is not an essential feature of social learning. What is crucial to social learning is merely the idea of the Internet providing access to information about the strategies, knowledge and experiences of human agents. There is no requirement here for real-time monitoring. Indeed, for the most part, we suspect that social learning will be undertaken with respect to previously acquired data sets, as opposed to real-time data streams. Such data sets include those that are already available. As is noted by Myaeng et al. [58], blog posts, online community services, and social media sites track the experiences and knowledge that humans have managed to distil from the environment and encode in digital form. Such resources provide the basis for what a Myaeng et al. [58] refer to as *experiential knowledge mining*, which is defined as "the process of acquiring experiential knowledge, as opposed to a priori knowledge, from a variety of multimedia sources that describe human experiences of various sorts" [58, p. 33].

From the standpoint of social learning, therefore, the human social environment serves as the target of a particular kind of knowledge acquisition. The main virtue of this vision is that it helps to expand the horizons of efforts that seek to emulate human-level competence in some domain of interest. In particular, it provides us with an alternative means of acquiring knowledge that might be difficult, costly or (perhaps) impossible to acquire via other means. Although this is clearly not the place to undertake a detailed analysis of the situations in which social learning is appropriate, we suspect that it is the nature of the target knowledge that determines the suitability of social learning approaches. In particular, we suggest that Internet-mediated social learning is perhaps best suited to the acquisition of knowledge that is tacit (i.e., difficult to verbalize), experiential (i.e., derived from experience and training), and socially-entrenched (i.e., socially-acquired and socially-manifested). The knowledge associated with the aforementioned ARV case possesses all these features. Such knowledge is, for the most part, tacit and therefore hard to express (at least in linguistic form). It is also knowledge that is acquired as a result of extensive experience and training and therefore qualifies as a form of experiential knowledge. Finally, it is a form of knowledge that is socially entrenched. This particular feature is difficult to describe in summary form, but it is perhaps most easily thought of as a form of knowledge that depends on the social environment for its acquisition or expression. When it comes to driving-related knowledge, for example, what we confront is a body of knowledge that is often tied to the interactions between individual drivers. Such knowledge is not the sort of knowledge that can be (easily) expressed independently of other social agents, and it is thus not the sort of knowledge that can be acquired (in a knowledge engineering sense, at least) without the support of a suitably rich social environment. It is this 'social' aspect that is perhaps most important when it comes to social learning. Tacit and experiential knowledge obviously present specific challenges to knowledge engineers; however, they are not beyond the reach of contemporary knowledge elicitation techniques (see [53]). Socially-entrenched knowledge is somewhat different, however, often requiring the observation and analysis of large-scale social interactions. It is in this sense, perhaps, that we can begin to understand the significance of the Internet from a knowledge engineering perspective [59]. By affording access to the human social environment, the Internet functions as a form of social observatory, enabling machines to prospect for epistemic gold in terrains that were previously beyond their reach.

B. Active Learning

We have seen how the Internet provides an unprecedented form of contact with the human social environment, opening up an array of opportunities for AI systems to observe and monitor human behavior. And we have seen how such forms of contact provide the basis for at least one form of machine learning. There is, however, a risk associated with this idea of the Internet functioning as a form of social observatory. The risk is that we lose sight of the way in which AI systems are able to play an active role in shaping the course of their own (socially-scaffolded) cognitive development. When we view the Internet as a form of observatory, there is a danger that we see machines as merely passive observers of the human social realm. This is, we suggest, a highly impoverished view of the learning opportunities made available by the Internet.

There is, however, no reason why we should restrict ourselves to this purely passive view of machine learning. There are a number of ways in which AI systems can play a more active role in socially-mediated learning processes. One example of this comes from studies into so-called citizen science systems [47]. One of the challenges confronting such systems is the need to ensure the continued engagement of the human community in the face of the human proclivity for boredom and distraction. A number of studies have attempted to address this problem by developing statistical models that predict the likelihood of user disengagement [60], and such models can be used to implement an array of intervention strategies that seek to sustain human interest [61][62]. As noted by Mao et al. [60]:

> The ability to predict forthcoming disengagement of individual workers would allow systems to make targeted interventions, such as providing especially interesting tasks to workers at risk of becoming bored, directing support to struggling new workers, helping with the timing of special auxiliary materials or rewards, and encouraging workers to return in the long run. (p. 1)

Needless to say, the ability to maintain user interest in a task is a relatively weak example of a machine playing an active role in socially-mediated learning. A better example comes from research into what is called *active learning* [63][64]. Active learning is a form of machine learning in which a machine learning system exerts some control over the learning process, actively structuring its training experiences in a manner that yields the best learning outcome. Such forms of control have been shown to yield a number of benefits. For example, active learning has been shown to improve the efficiency of the learning process by reducing the number of training examples that are required to reach near-optimal levels of performance on an image processing task [65].

For the most part, active learning involves the adaptive selection and sequencing of specific training experiences. In the context of an image processing task, for example, an active learning system may decide what images will be the focus of learning efforts, as well as the order in which the images will be processed. Such decisions are typically informed by routines that estimate the optimality of different response options relative to the system's current state, previous learning experiences, and overall learning objectives. In this sense, active learning systems can be seen to implement something akin to a 'metacognitive' ability, with one form of 'cognitive' processing (i.e., that associated with optimality assessments) influencing the behavior of other parts of the cognitive economy (e.g., the shape of specific learning routines).

A good example of active learning in an Internet context is provided by Barrington et al. [66]. Barrington et al. describe the use of an online game, called Herd It, in which groups of human individuals are required to annotate a musical resource with descriptive tags. These annotations are used to train a supervised machine learning system that ultimately aims to perform the annotation task independently of the human agents. All of this is broadly in line with the general shape of machine learning; but what makes Barrington et al.'s system of particular interest is the way in which the machine shapes the course of its own learning by actively selecting the resources to be annotated by human players. This is important, because it gives the machine an opportunity to select those forms of feedback that are likely to be of greatest value relative to its subsequent 'cognitive' development. In the words of Barrington et al. [66], "the machine learning system actively directs the annotation games to collect new data that will most benefit future model iterations" (p. 6411).

A consideration of active learning thus expands our understanding of the forms of contact that the Internet provides with the human social environment. Rather than seeing the Internet solely as a form of social observatory—one that permits a largely passive form of observational contact with humanity—we can now entertain a more active (and interactive) view of the Internet. On this view, the Internet provides machines with an opportunity to influence human behavior, altering the nature of the information flows that underpin the emergence of specific forms of cognitive proficiency.

C. Language Learning

The advent of the Internet (and especially the Web) has led to a burgeoning of research interest into all things linguistic. Such interest is evidenced by research into Natural Language Processing (NLP) (e.g., [67]), information extraction [68], and sentiment analysis [69]. Other research efforts aim to develop various forms of language-enabled agents, i.e., computational agents that are able to exhibit proficiency in the use of natural language expressions. Work in this area includes research into so-called social bots [70], chatbots [71] and conversational agents [72].

The reason for this renewed interest in language-related technologies is, at least in part, due to the wealth of linguistic content that is available in the online realm. Such content provides us with a substantive body of linguistic data that can be used to inform large-scale analytic efforts. It should also be clear that the Internet has transformed the incentives that drive research and development in this area—consider, for example, the use of Twitter feeds as a means of predicting (and perhaps influencing!) the outcome of political elections [73]. The upshot is that language learning has become an important focus of attention for the machine learning community.

How does this renewed interest in linguistic analysis impact the present discussion on machine intelligence? The most obvious answer to this question is that machines will become increasingly proficient in understanding human language, and as a result of this understanding, they will be better placed to exploit our linguaform contributions to the online realm (e.g., they will have an improved ability to distil information and knowledge from resources such as Wikipedia, Twitter, Facebook and so on). It should also be clear that enhancements in linguistic proficiency often go hand-in-hand with improvements in communicative ability. There can be little doubt that such communicative abilities play an important role in extending the cognitive reach of an agent community. Indeed, we might be inclined to view communication as a form of networking capability that enables agents to 'connect' with an array of cognitively-potent resources. This applies as much to human agents as it does to their synthetic counterparts. As noted by Merlin Donald [74], when it comes to human language, "[i]ndividuals in possession of reading, writing, and other visuographic skills...become somewhat like computers with networking capabilities; they are equipped to interface, to plug into whatever network becomes available" (p. 311). Linguistic competence can therefore be seen to work in concert with other forms of scaffolded development, enabling machines to distil knowledge from online textual sources and providing the basis for communicative exchanges with human agents. Such capacities are likely to be of crucial importance when it comes to the social scaffolding of machine intelligence.

Communication is no doubt important when it comes to the ability of machines to press maximal cognitive benefit from the human social environment. But there are other ways to think about the cognitive significance of language. Of particular interest is what is sometimes called the supracommunicative view of language function [75][76]. The general idea, in this case, is that language plays a role in enhancing, transforming, or otherwise altering the cognitive capabilities of the language-wielding agent. There are a number of ways of unpacking this claim; for present purposes, however, we will limit our attention to three (not altogether distinct) manifestations of the supracommunicative view. These are the transformative, the augmentative and the configurative/programmatic views.

The transformative view derives from the work of the philosopher, Daniel Dennett [77]. Dennett suggests that our ontogenetic immersion in a linguistic environment contributes to an effective reorganization of the human cognitive economy, yielding a shift from parallel processing into something that more closely resembles the information processing profile of a conventional (symbol-manipulating) computational machine. Interestingly, Dennett proposes that some of the most distinctive features of human cognition (including human consciousness) emerge as a result of our attempts to get to grips with the linguistic domain. Inasmuch as we accept these claims, it should be clear that a simple communicative view of language is unlikely to do justice to the potential impact of the Internet on future forms of machine intelligence: By immersing intelligent systems in a linguistically-rich environment, and by forcing such systems to assimilate linguaform representations deep into their cognitive processing routines, we potentially endow machines with the sorts of abilities and insights that only us language-wielding human agents are able to grasp.

Another take on the cognitive role of language comes in the form of the augmentative view. The most vocal proponent of this view is Andy Clark [75][76][78]. Clark see language as a particularly potent form of socially-derived cognitive scaffolding that performs a variety of cognition-enhancing roles:

> Embodied agents encounter language first and foremost as new layers of material structure in an already complex world. They also come to produce such structures for themselves, not just for communicative effect but as parts of self-stimulating cycles that scaffold their own behaviour. These layers of structure play a variety of cognition-enhancing roles. They act as new, perceptually simple targets that augment the learning environment, they mediate recall and help distribute attention, they provide a key resource for freezing and inspecting complex thoughts and ideas, and they seem fit to participate in truly hybrid representational ensembles. All these benefits are available both 'online' (in the presence of written words on a page, or sounds in the air) and then 'offline' (thanks to covert self-stimulating cycles that engage much of the same machinery used in the ecologically primary case). [79, p. 373]

Empirical support for the augmentative view comes from a variety of quarters [39][80][81]. In studies with human subjects, language has been shown to play a productive role in category learning [39][81], and such effects have also been observed in computer simulations, with linguistic labels supporting category learning in artificial neural networks [82][83]. Exposure to a linguistic environment also appears to bolster the cognitive performance of certain non-human animal species, such as chimpanzees [84] and parrots [85]. Although these studies are often seen as failures from a language learning perspective (no one doubts, for instance, that the animals in these studies failed to acquire human-level language abilities), the studies are nevertheless remarkable in demonstrating that even minimal forms of linguistic competence are able to augment the cognitive profile of such animals, and they do so in ways that are reminiscent of human-level cognitive achievements [84].

A final way to unpack the supracommuncative view of language is to emphasize the way in which language can be used to configure and control a set of cognitively-relevant resources. This is what we will call the configurative/programmatic view of language. Perhaps the most explicit expression of the idea behind this view is provided by Lupyan and Bergen [81]. They see language as a tool or control system that can be used to 'program' the mind. In the case of human minds, for example, Lupyan and Bergen [81] highlight the ways in which linguistic stimuli can be used to shape aspects of the human cognitive economy, presumably by altering the dynamics of neural processing. Exposure to linguistic stimuli has thus been shown to alter certain forms of perceptual processing, boosting the extent to which previously unseen objects enter visual awareness [86]. Linguistic cues can also be used to activate and reactivate certain forms of mental content. The activation of visual images, for example, typically depends on visual input. With language, however, we are able to exert control over our imaginative faculties. A mental image of the Colosseum, for example, can be evoked simply by exposing our minds to the word "Colosseum" (see [39]).

To some extent the configurative/programmatic view bears much in common with the transformative and augmentative views of language. Clark, for example, has often appealed to the idea of language as a tool that helps to tame the restive information processing dynamics of the biological brain. "Encounters with words and with structured linguistic encodings," he suggests, "act so as to anchor and discipline intrinsically fluid and context-sensitive modes of thought and reason" [87, p. 263]. A key difference between the augmentative and configurative/programmatic views emerges in respect of the nature of the resources that are controlled by linguistic stimuli. In the case of the configurative/programmatic view, it is not just our own minds that are controlled via language, it is also the minds of others. And in shaping the minds of others, we are able to exert some degree of control over the social environment:

> We can sculpt the minds of others into arbitrary configurations through a set of instructions, without having to go through laborious trial-and-error learning. We can cause someone to imagine something, to recall a memory, to do (or not do) something. [81, p. 409]

The result is a view of language as a form of generic control system, one that can be used to configure (and thereby shape the behavior of) a variety of disparate resources. When applied to the social domain, the configurative/programmatic view helps us see language as on a par with physical action, in that it can be used to intervene in the social environment (just as physical actions can be used to alter the structure of the physical environment). This is an image that dovetails with the earlier discussion of active learning (see Section V-B). For in using language to manipulate the minds of others, there can be little doubt that we are in possession of a tool for structuring the nature of our contact with the human social world. In this sense, language affords a degree of control over socially-scaffolded forms of development.

The communicative, transformative, augmentative and configurative/programmatic views thus provide us with a complex picture of the cognitive impact of language. In directing their learning efforts to the linguistic realm, machines are potentially poised to exploit some of the cognitive virtues of language. Such virtues are perhaps most easily understood with respect to the augmentative and transformative views; however, in developing linguistic competence, we should not forget that language also influences the nature of the cognitive contact that machines have with the human social environment, opening up new arenas for scaffolded development (the communicative view) and providing new opportunities for machines to shape the structure of their learning experiences. Inasmuch as we aspire to build machines that emulate the performance profile of the human cognitive system, language learning thus looks to be of crucial importance. It may indeed be the case that human-level cognizing is inextricably linked to language, and that an ability to emulate human cognition is predicated on an ability to negotiate the linguistic domain. Something along these lines is, in fact, suggested by Mirolli and Parisi [39]. Commenting on the role of language in the development of robotic systems, they suggest that it "may be impossible to develop a human-like cognitive robotics without endowing robots with the capacity of using language for themselves as humans do" (p. 301).

D. Predictive Learning

According to an increasingly popular theory in theoretical neuroscience, the biological brain is a hierarchically-organized system in which higher-level neural regions are engaged in a continuous effort to the predict the activity of lower-level neural regions [88][89]. This model—which we will dub the predictive processing model—has proved attractive for a variety of reasons. For example, it provides an explanation for reciprocal connectivity between anatomical brain regions, and it also promises a unified account of perception, action and cognition [89][90]. The model has also proved attractive with respect to the recent efflorescence of research into deep learning [91]. Deep learning systems thus incorporate some of the features of the predictive processing model, and this may account for their superior performance in a number of task domains.

The kind of learning implemented by the predictive processing model of brain function is perhaps best characterized as *predictive learning*. It is a form of learning in which higher-levels of the predictive processing hierarchy seek to predict the activity of lower levels. In the brain, this learning is assumed to be driven by prediction error, reflecting the mismatch between predicted and actual patterns of brain activity. In essence, the goal of predictive learning is to minimize the global prediction error that is generated by the biological brain as part of its attempt to predict what is in effect its own activity. Given that such activity is ultimately tied to the external environment (via the receipt of sensory information), predictive learning leads to structural changes that reflect the brain's attempt to secure a predictive grasp of the environment in which it is embedded. This is important, for it is believed that one of the outcomes of predictive learning is the establishment of a generative model that reflects the causal structure of the local learning environment. “A generative model,” Clark [89] suggests, “...aims to capture the statistical structure of some set of observed inputs by inferring a causal matrix able to give rise to that very structure” (p. 41). It is in this sense that predictive learning is sometimes seen to yield models that embody the causal structure of mechanisms that give rise to bodies of sensory information:

> In brief, biological systems can distil structural regularities from environmental fluctuations (like changing concentrations of chemical attractants or sensory signals) and embody them in their form and internal dynamics. In essence, they become models of causal structure in their local environment, enabling them to predict what will happen next and counter surprising violations of those predictions. [92, p. 2101]

It is here that we begin to creep up on a novel, albeit contentious, proposal regarding the role of the Internet in supporting the emergence of advanced forms of machine intelligence. In short, the idea is that in the attempt to form a generative model of data that derives from the human social environment, a hierarchically-organized predictive processing system may come to acquire a ‘deep understanding’ of human behavior at both the individual and collective (social) levels. This ‘deep understanding’ is reflected in the way in which a predictive processing system comes to embody the causal structure of the social domain. A generative model of the human social environment can thus be seen to lead to a deep understanding of the causal processes that govern the shape of human behavior, just as the operation of brain-based predictive processing regimes are deemed to yield a deep understanding of the causal processes that govern the structure of incoming sensory information [90]. A good probabilistic generative model for individual human behavior (or larger-scale patterns of social flux) would therefore seek to capture the ways that patterns of human behavior are generated by an inferred nexus of interacting distal causes.

This idea is suggestive, for it may help to shed light on the mechanisms that underlie various forms of social intelligence. When it comes to the realms of individual human behavior, for example, the notion of acquired generative models may help us understand the basis for folk psychological characterizations of the behavior of both ourselves and others. Our conventional approach to explaining human behavior in terms of beliefs, hopes, fears, desires and dreams may thus reflect nothing more than our attempt to gain a predictively- and explanatory-potent toehold on the social realm, with human psychological states being ascribed to individual agents as part of the brain’s attempt to make sense of complex bodies of social data. Such a view may provide insight into some of the most mysterious elements of our mental lives, including that ever-elusive phenomenon we call conscious experience. Perhaps, for example, we are aware of ourselves as psychological agents precisely because we model our own behavior in the way we model others. If true, the result is a view of human consciousness that appeals to the way in which the shape of our own mental lives owes much to the structure of the social environment in which we are embedded. In essence, the idea is that we should understand human consciousness as a form of socially-scaffolded cognition.

Echoes of this sort of view can, in fact, be found in the works of Lev Vygotsky, one of the pioneers of scaffolded cognition research:

> The mechanism of social behavior and the mechanism of consciousness are the same. We are aware of ourselves in that we are aware of others; and in an analogous manner, we are aware of others because in our relationship to ourselves we are the same as others in their relationship to us. [93, p. 29]

The view is also evident in work of a more recent nature. Graziano and Kastner [94], for example, describe a theoretical account of self-awareness that is rooted in an appeal to socially-oriented predictive processes:

> In the present hypothesis, the human brain evolved mechanisms for social perception, a type of perception that allows for predictive modeling of the behavior of complex, brain-controlled agents. There is no assumption here about whether perception of others or perception of oneself emerged first. Presumably they evolved at the same time. Whether social perception is applied to oneself or to someone else, it serves the adaptive function of a prediction engine for human behavior. [94, p. 109]

Inasmuch as such accounts provide insight into the forces and factors that give rise to human conscious experience, they may help to reveal the significance of socially-oriented predictive learning to the creation of conscious machines.

This is, to be sure, a grand claim, and no doubt many issues need to be resolved before the idea can be taken seriously. One of these relates to the nature of the informational contact that machines have with the human social environment. Do the digital traces provided by the online realm provide us with a sufficiently rich and detailed representation of the dynamical profile of human behavior, one that is apt to yield (via predictive learning) a generative model that traces the causal contours of human behavior at both the individual and collective levels? The answer to this question is unclear at the present time, although it should be noted that the Internet plays an increasingly important role in a variety of human activities, and it is thus poised to provide ever-more detailed insights into the shape of human behavior. Crucially, the success of some predictive analytics platforms already attests to the predictive potential of at least some forms of online data. New predictive apps, such as Google Now, for example, are able to make predictions based on the analysis of various data sources, and they do so in a way that is sometimes seen to belie an uncanny knowledge of their user base.

This is not to say that the view of the human social environment as provided by the Internet will be exactly the same as that enjoyed by a human individual. There are clearly important differences in the kind of information that is accessible to an online machine learner as opposed to the information that is made available to a human observer of the social realm. It is not clear, however, that such differences should always be seen as placing machine-based systems at a disadvantage. Consider, for instance, the way in which the Internet affords a panoptic view of social processes that operate at a variety of social scales (e.g., at the level of teams, groups, organizations, communities and societies). The application of predictive learning to such bodies of social data may give rise to generative models that embody the causal structure of social mechanisms (i.e., the mechanisms that govern the behavior of social systems). In essence, we suggest that in the attempt to secure a predictive grasp on bodies of social data, a machine learning system could be forced to approach a social system in much the same way that it approaches an individual human agent, yielding a generative model that captures some of the hidden causal forces that operate (perhaps) exclusively at the social or societal level. The result is a rather unique vision of machine intelligence. It is a vision in which social systems are themselves perceived as psychologically-rich and complex entities. And it is a vision in which the goal of learning is to make sense of the social world—to develop a deep understanding of the various forces and factors that govern the flux of social data. A social machine, it seems, is not just a machine that is situated or embedded within society (although that is indeed the case). Neither is it simply a machine whose 'mind' is, in some sense, a product of society (although that it is also true). A social machine is a machine that is, we suggest, poised to develop a novel kind of mind, a mind that is specifically oriented to the social realm—a mind of society.

Some insight into the potential value of socially-oriented predictive learning is provided by recent studies using deep learning techniques. One such study is described by Phan et al. [95]. They used a combination of computational ontologies and deep learning techniques to yield a system that generated predictions of individual human behavior in the health domain. What is interesting about this study is that by incorporating structured representations of domain-specific knowledge (in the form of ontologies) into the learning regime, the resultant system was able to not only predict human behavior, but also generate explanations for such behavior. Such results, Phan et al. [95] suggest, indicate a "deep understanding of...human behavior determinants" (p. 311).

Another interesting application of deep learning techniques concerns the attempt by Vondrick et al. [96] to predict human action sequences from video images. This study is of particular interest because the training corpus for the deep learning system consisted of 600 hours of unlabeled video downloaded from the YouTube website. Vondrick et al.'s study thus exemplifies one of the ways in which the Internet/Web provides a form of informational access to the human social environment in a manner that can be used to support the development of predictive capabilities. YouTube videos are, of course, uploaded by human users, and they do not always afford an unfiltered insight into what we might call ecologically-normal patterns of human behavior. The step from YouTube to more direct and real-time observational data streams is, however, a short one. There is no reason, for example, why the approach of Vondrick et al. [96] could not be applied to the data provided by Internet-enabled video recording devices, such as webcams and CCTV devices.

Finally, consider a study by Lv et al. [97] involving the use of deep learning methods for the purposes of traffic flow prediction. This study is interesting for a variety of reasons. Firstly, Lv et al. remind us of the wealth and diversity of information that can be used for predictive purposes. This includes information from "inductive loops, radars, cameras, mobile Global Positioning System, crowd sourcing, social media, etc." (p. 865). A second point of interest concerns the focus of Lv et al.'s study, which is nicely aligned with the earlier discussion of social learning (recall the discussion of the ARV case in Section V-A). This is a useful reminder that one form of learning (e.g., predictive learning) can be used to support other forms of learning (e.g., social learning). Finally, note that the target of Lv et al.'s [97] study is a 'collective system' comprised of multiple elements (i.e., vehicles), each of which is controlled by a human agent. This is, as such, a nice example of the attempt to model the behavioral profile of a particular form of 'social system'. In this respect, Lv et al.'s study provides some insight into the sorts of approaches that might be relevant to the acquisition of socially-oriented generative models.

E. *Incremental Learning*

The notion of socially-scaffolded cognition encourages us to take a developmental perspective with respect to machine intelligence. In particular, we are encouraged to see machine-based cognitive capabilities as emerging from a developmental matrix that includes (among other things) the human social environment. In considering the opportunities for socially-scaffolded development, however, it is easy to overlook the fact that the cognitive wherewithal of human infants is not the same as their adult counterparts. It is here that we encounter a productive point of contact with work that shows how maturational shifts in cognitive, sensory and motor capabilities may be of crucial relevance to the emergence of advanced forms of cognitive competence [98]–[103].

The idea that 'immaturity' may be of adaptive value with regard to the ontogenetic emergence of certain capabilities was first discussed by Turkewitz and Kenny [104]. According to their hypothesis, immaturity alters the kind of information a learning system can process, thereby altering what is sometimes called the 'effective' structure of the learning environment. During the initial stages of development, the complexity of the learning environment is reduced as a result of the relative immaturity of sensorimotor systems. As development proceeds, however, maturational processes lead to the progressive attenuation of initial processing constraints, limitations, and biases, and this, in turn, leads to an increase in the complexity of the training data. When all of this is applied to the cognitive domain, the result is a proposal regarding the role of maturational parameters in the acquisition of advanced forms of cognitive competence. According to this proposal, various forms of 'immaturity' may be of crucial significance when it comes to a cognitive system's ability to achieve the sorts of cognitive success that mark the end of the developmental process.

What are the implications of this proposal for socially-scaffolded forms of machine intelligence? Perhaps the best answer to this question comes from research into a specific form of machine learning, known as *incremental learning* [51]. Incremental learning, as defined by Kirby and Hurford [105], is:

> ...the idea of some learning-related resource starting at a low value, which then gradually increases while (but not necessarily because) the organism matures. Also essential to incremental learning is the proposition that the initial low (immature) value of the resource actually facilitates, or even enables, the early stages

> of learning. Later stages of learning are in turn facilitated, or enabled, by higher-valued settings of the resource concerned. [105, p. 4]

Some insight into the potential importance of incremental learning is revealed by a classic study by Jeffrey Elman [99]. Elman sought to determine whether a particular kind of artificial neural network, called a recurrent neural network, could acquire a form of grammatical competence characterized by an ability to learn about verb agreement and clause embedding in sentences such as: "The girls who the teacher has picked for the play which will be produced next month practice every afternoon" [99, p. 4]. As part of the training regime, the sentences were presented to the network one word at a time, and the main objective of the network was to predict the next word in the sentence. As Elman [99] notes, this task "forces the network to develop internal representations which encode the relevant grammatical information" (p.5).

Unfortunately, Elman's initial attempts to get the network to learn about grammatical structure were in vain. Not only did the network fail to develop a fully generalizable performance profile, it also failed to adequately master the data on which it was trained. As part of the effort to account for these results, Elman deployed an alternative training regime, one in which the network was initially presented with examples of very simple sentences and then progressively exposed to the more complex ones. The aim was to isolate the precise point at which the network's performance broke down—at what level of sentential complexity would the network prove to be incapable of making further progress?

The results of this alternative training regime were surprising. Elman discovered that when presented with staged training inputs (each increasing in complexity) the network was able to realize its original training objectives. Indeed, what seemed to be important to the network's ultimate ability to learn about grammatically complex sentences was that its training regime was structured in such a way that it was able to learn about the simple cases first. Once the network was proficient in handling these simple cases, it was then able to deal with the more complex cases. It was almost as if the network's success with the simple cases laid the foundation for subsequent success in dealing with the more complex cases.

Moving on from this result, Elman explored the effect of a further manipulation. In this case, rather than impose restrictions on the sequential complexity of the training inputs, Elman used an incremental memory solution in which the recurrent feedback (provided by a layer of context units) was gradually increased as training progressed. The effect of this manipulation was to limit the temporal window in which linguistic inputs could be processed, thereby forcing the network to focus (at least initially) on the simplest training cases. Then, as the memory provided by the recurrent units was increased over the course of training, the network was able to deal with progressively more complex inputs. The effect of the incremental memory solution was thus the same as that achieved by the staged input training case: it promoted an initial under-sampling of the training data in such a way that the network's long-term ability to learn about complex grammatical regularities was enhanced.

As Elman notes, this is an important discovery, because it may help to shed light on the functional significance of a developmental progression in neurocognitive resources. Thus, rather than see the working memory limitations of young children as a computational shortcoming that needs to be overcome in order to reveal the functional profile of adult cognition, Elman's findings suggest that immature cognitive capabilities may play an important (and perhaps indispensable) role in enabling young infant minds to acquire adult forms of linguistic competence. Commenting on the role of memory limitations in language learning, Elman states that:

> ...the early limitations on memory capacity assume a more positive character. One might have predicted that the more powerful the network, the greater its ability to learn a complex domain. However, this appears not always to be the case. If the domain is of sufficient complexity, and if there are abundant false solutions, then the opportunities for failure are great. What is required is some way to artificially constrain the solution space to just that region which contains the true solution. The initial memory limitations fulfil this role; they act as a filter on the input, and focus learning on just that subset of facts which lay the foundation for future success. [99, p. 9–10]

Here, then, is one example where a form of limited or restricted processing may help an agent achieve success in what could otherwise prove to be an intractable problem domain. By imposing a set of constraints on the kinds of information structures that can be processed, maturational processes can be seen to support the progressive reshaping of the effective structure of the linguistic environment, or at least the nature of the language learning task that confronts the learning system. Perhaps this insight goes some way toward understanding the problems that adult humans often experience in learning a second language [106][107]. In learning a new language, it might be thought that adults are in a much better position than young infants. And, at least during the early stages of language learning, adults do indeed appear to make more progress [100]. Their early successes, however, appear to come at a substantial cost: as time passes, the young infants quickly overtake the adult learners and rapidly become proficient in the target language. In this case, the early constraints in cognitive processing seem to be playing a productive role in enabling the human infant to approach the language learning task in the most effective manner.

Language is not the only domain where developmentally-significant alterations in maturational parameters have been studied in a machine learning context. An important source of information regarding the functional role of early limitations in the development of advanced cognitive and behavioral capabilities comes from recent work in developmental and evolutionary robotics [98][102]. Gómez et al. [98], for example, describe an intriguing set of results pertaining to the development of sensorimotor capabilities in a real-world robotic system. They report that a developmental profile characterized by progressive increments in the complexity of sensory, motor and neurocomputational subsystems results in a profile of task performance that is superior to that of a robot in which the relevant maturational processes are disabled. Commenting on this developmentally-grounded dissociation in 'adult' performance profiles, they suggest that:

> ...rather than being a problem, early morphological and cognitive limitations effectively decrease the

> amount of information that infants have to deal with, and may lead to an increase in the overall adaptivity of the organism. [98, p. 119]

More recent studies, again using real-world robots, have extended these results to the domain of social cognition. Nagai et al. [108] thus demonstrate that gradual increments in the spatiotemporal resolution of a robot's visual system enables it to discriminate between actions that are generated by itself and other agents. Immature vision, Nagai et al. [108] suggest, helps to shape the early perceptual environment of a system in such a way as to support the subsequent emergence of socially-relevant capabilities, such as the ability to discriminate the 'self' from others and engage in imitative behavior.

The general lesson to emerge from research on incremental learning is that early limitations in one or more parameters of a cognitive system (human or machine) may play a productive role in enabling that system to deal with the challenges of a complex (and perhaps otherwise intractable) problem domain. This seems to be of particular importance when it comes to the sorts of challenges faced by socially-situated machines. For such systems must learn to negotiate a highly complex domain, characterized by linguistic expressions and digital traces of human behavior. In tackling such domains, it may be necessary to recapitulate some of the maturational processes that operate in the case of human cognitive development. We have already seen how this sort of idea might be applied to the realm of language learning—recall the work by Elman [99]—and extensions of this work may be relevant to the attempt to furnish machines with more advanced natural language processing abilities (see Section V-C). Incremental learning may also be important when it comes to the attempt to develop predictive models of human behavior (see Section V-D) or the attempt to press maximal cognitive benefit from various forms of social learning (see Section V-A). In all these cases, the target domain concerns some aspect of human behavior, and the objective of the learner is to achieve the sort of competence that enables them to navigate, explore and negotiate the complexities of the human social world. In the human case, it looks likely that issues of development and maturation play a potentially crucial role in enabling infant minds to develop into fully-fledged adult cognizers. Inasmuch as this is true, is there any reason to think that AI systems will be able to bypass a stage of relative cognitive immaturity? One of the goals of AI is to implement systems that exhibit the capabilities characteristic of human *adults*. But inasmuch as human cognitive capabilities emerge as the result of maturation and socially-scaffolded development, is there any reason to think that AI systems can forgo the equivalent of a larval stage and proceed directly to the end-stage of the cognitive developmental process? Such an assumption looks to be particularly precarious if we accept that the human social environment forms part of a complex developmental system that drives human forms of mental metamorphosis.

VI. Conclusion and Future Work

There are good reasons to think that our status as social animals and our embedding within a social environment are of crucial importance when it comes to understanding the distinctive features of the human cognitive system. This is the case, irrespective of whether our attention is focused on issues of phylogeny or ontogeny. From a phylogenetic perspective, the human mind may have evolved to deal with the challenges and demands of a social environment whose complexity and variability increased across the course of human evolution. Similarly, from an ontogenetic perspective, it seems that the human social environment may have played a crucial role in shaping the course of cognitive development, enabling a cognitively altricial human infant to emerge as one of Planet Earth's most precocious cognizers. This is, to be sure, a compelling image. It is an image in which the human mind is viewed as a product of the human social environment, a device of our own creation, a socially-manufactured cognitive machine.

But it is not just our view of human intelligence that stands to be transformed by this image; it is also our view of machine intelligence. For inasmuch as we strive to build AI systems in our own cognitive image—as machines that emulate our own, species-specific form of intelligence—then it is surely worth considering the extent to which the human social environment is poised to play a productive role in yielding the next generation of intelligent machines.

This is the idea that we have attempted to develop in the present paper. Our claim is that the Internet provides an unprecedented form of informational contact with the human social environment, and that this contact occurs at multiple levels of social organization, from individual human agents through to teams, groups, communities, and societies. By virtue of this contact, the Internet enables AI systems to be the beneficiaries of various forms of socially-scaffolded cognitive development. In short, we suggest that the Internet provides opportunities for the implementation of what we called social machines, i.e., machines that are able to benefit (in a cognitive sense) from their contact with humanity. Thus construed, a social machine is similar to a socially-situated robot [38][41]. The main difference is that a social machine operates in the online realm, and the limits of its social ecology are thus co-extensive with the social reach of the Internet.

Needless to say, there are various ways in which the present work could be extended (an no doubt improved). One area for future work concerns the kinds of systems that are able to benefit from their contact with the human social environment. In particular, it is unclear whether a particular kind of computational substructure—such as a hierarchically-organized predictive processing economy—is a prerequisite for human-like forms of cognitive competence. In addition to research into machine learning, therefore, future work should aim to consider the kinds of cognitive architecture that may be required to support socially-scaffolded development.

Another area of interest concerns the relevance of additional forms of learning. In addition to the forms of learning discussed in the present paper (i.e., social, active, language, predictive, and incremental learning), it may be important to consider learning mechanisms that bias, direct or promote interest in the social environment. It seems likely that an ability to benefit from social-scaffolding, and indeed the status of a system as a socially-situated agent, may depend on the sensitization of reward mechanisms to social feedback, or the implementation of motivational mechanisms that encourage or enable socially-oriented forms of learning. In this respect, it may be useful to consider work relating to reinforcement learning, intrinsic motivation and curiosity-driven learning [109]–[111]. There is, of course, no reason why these forms of learning (as well as those discussed in the present paper) should be studied independently, and there is likely to be considerable merit in combining different forms of learning within a single system.

A further area of work relates to the application of the present approach to other kinds of cognition. For the most part, we have limited our attention to scaffolded cognition. In future work, it may be useful to extend this analysis to other cognitive kinds, such as extended, embedded, and embodied cognition. (See Smart [11][112] for some initial steps in this direction.)

The cognitive significance of the Internet is typically judged relative to its impacts on human cognition [113][114]. This is, of course, understandable. It is natural for us to wonder (and sometimes worry) about the implications of the Internet for our species, especially when it comes to its effects on our cognitive capabilities. For such capabilities are the hallmark of humanity: it is our cognitive profile that sets us apart from other forms of terrestrial life, and it is such capabilities that enable us (and only us) to actively shape the course of our cognitive destiny—to engineer something like the Internet and then worry about its cognitive consequences. But the cognitive implications of the Internet do not end at the borders of the human mind. In creating the Internet, our species has established a new kind of informational ecology, one that opens up new opportunities for research into machine intelligence. In the present paper, we have focused on one particular opportunity. We have suggested that the Internet provides an unprecedented form of informational contact with the human social environment, enabling machines to exploit opportunities for socially scaffolded cognitive development. It is through such forms of contact, perhaps, that we will witness the emergence of a new kind of cognitive machine, a machine whose mind is as much a product of society as are the human minds it seeks to emulate.

ACKNOWLEDGMENT

This work is supported under SOCIAM: The Theory and Practice of Social Machines. The SOCIAM Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1 and comprises the Universities of Southampton, Oxford and Edinburgh.

REFERENCES

[1] P. R. Smart, "Machine intelligence and the social web: How to get a cognitive upgrade," in *9th International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE'17)*, V. Gripon, O. Chernavskaya, P. R. Smart, and T. T. Primo, Eds. Athens, Greece: International Academy, Research, and Industry Association (IARIA), 2017, pp. 96–103.

[2] J. Sutton, "Scaffolding memory: Themes, taxonomies, puzzles," in *Contextualizing Human Memory*, C. Stone and L. Bietti, Eds. New York, New York, USA: Routledge, 2016.

[3] K. Sterelny, "Minds: Extended or scaffolded?" *Phenomenology and the Cognitive Sciences*, vol. 9, no. 4, pp. 465–481, 2010.

[4] A. Clark and D. Chalmers, "The extended mind," *Analysis*, vol. 58, no. 1, pp. 7–19, 1998.

[5] A. Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York, New York, USA: Oxford University Press, 2008.

[6] R. Menary, Ed., *The Extended Mind*. Cambridge, Massachusetts, USA: MIT Press, 2010.

[7] R. D. Rupert, "Challenges to the hypothesis of extended cognition," *Journal of Philosophy*, vol. 101, no. 8, pp. 389–428, 2004.

[8] P. Robbins and M. Aydede, Eds., *The Cambridge Handbook of Situated Cognition*. New York, New York, USA: Cambridge University Press, 2009.

[9] E. Hutchins, *Cognition in the Wild*. Cambridge, Massachusetts, USA: MIT Press, 1995.

[10] L. A. Shapiro, Ed., *The Routledge Handbook of Embodied Cognition*. New York, New York, USA: Routledge, 2014.

[11] P. R. Smart, "Situating machine intelligence within the cognitive ecology of the Internet," *Minds and Machines*, vol. 27, no. 2, pp. 357–380, 2017.

[12] A. Clark, *Being There: Putting Brain, Body and World Together Again*. Cambridge, Massachusetts, USA: MIT Press, 1997.

[13] E. Arnau, S. Ayala, and T. Sturm, "Cognitive externalism meets bounded rationality," *Philosophical Psychology*, vol. 27, no. 1, pp. 50–64, 2014.

[14] C. Craver, "Constitutive explanatory relevance," *Journal of Philosophical Research*, vol. 32, pp. 3–20, 2007.

[15] P. Ylikoski, "Causal and constitutive explanation compared," *Erkenntnis*, vol. 78, no. 2, pp. 277–297, 2013.

[16] S. O. Palermos, "Loops, constitution, and cognitive extension," *Cognitive Systems Research*, vol. 27, pp. 25–41, 2014.

[17] C. F. Craver and L. Darden, *In Search of Mechanisms: Discoveries Across the Life Sciences*. Chicago, Illinois, USA: The University of Chicago Press, 2013.

[18] D. M. Kaplan, "How to demarcate the boundaries of cognition," *Biology & Philosophy*, vol. 27, no. 4, pp. 545–570, 2012.

[19] V.-P. Parkkinen, "Developmental explanation," in *New Directions in the Philosophy of Science*, M. C. Galavotti, D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, and M. Weber, Eds. London, UK: Springer, 2014.

[20] D. Little, "Disaggregating historical explanation: The move to social mechanisms," in *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, S. Glennan and P. M. Illari, Eds. New York, New York, USA: Routledge, 2018.

[21] A. C. Love, "Developmental mechanisms," in *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, S. Glennan and P. M. Illari, Eds. New York, New York, USA: Routledge, 2018.

[22] R. I. M. Dunbar, "The social brain hypothesis," *Evolutionary Anthropology*, vol. 6, no. 5, pp. 178–190, 1998.

[23] R. W. Byrne and A. Whiten, *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford, UK: Oxford University Press, 1988.

[24] E. Herrmann, J. Call, M. V. Hernández-Lloreda, B. Hare, and M. Tomasello, "Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis," *Science*, vol. 317, no. 5843, pp. 1360–1366, 2007.

[25] H. Kummer, L. Daston, G. Gigerenzer, and J. B. Silk, "The social intelligence hypothesis," in *Human By Nature: Between Biology and the Social Sciences*, P. Weingart, S. D. Mitchell, P. J. Richerson, and S. Maasen, Eds. Mahwah, New Jersey, USA: Lawrence Erlbaum Associates, 1997.

[26] G. Miller, *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature*. London, UK: Vintage, 2000.

[27] M. V. Flinn, D. C. Geary, and C. V. Ward, "Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence," *Evolution and Human Behavior*, vol. 26, no. 1, pp. 10–46, 2005.

[28] M. V. Flinn, "Culture and developmental plasticity: Evolution of the social brain," in *Evolutionary Perspectives on Human Development*, 2nd ed., R. L. Burgess and K. MacDonald, Eds. Thousand Oaks, California, USA: Sage Publications, 2005.

[29] M. Ridley, *The Red Queen: Sex and the Evolution of Human Nature*. New York, New York, USA: Perennial, 2003.

[30] N. K. Humphrey, "The social function of intellect," in *Growing Points in Ethology*, P. P. G. Bateson and R. A. Hinde, Eds. Cambridge, UK: Cambridge University Press, 1976.

[31] E. Jablonka, S. Ginsburg, and D. Dor, "The co-evolution of language and emotions," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1599, pp. 2152–2159, 2012.

[32] Z. Petanjek *et al.*, "Extraordinary neoteny of synaptic spines in the human prefrontal cortex," *Proceedings of the National Academy of Sciences*, vol. 108, no. 32, pp. 13 281–13 286, 2011.

[33] K. Sterelny, "An alternative evolutionary psychology?" in *The Evolution of Mind: Fundamental Questions and Controversies*, S. W. Gangestad and J. A. Simpson, Eds. New York, New York, USA: The Guilford Press, 2007.

[34] L. Malafouris, *How Things Shape the Mind: A Theory of Material Engagement*. Cambridge, Massachusetts, USA: MIT Press, 2013.

[35] M. V. Flinn and K. Coe, "The linked red queens of human cognition, coalitions, and culture," in *The Evolution of Mind: Fundamental Questions and Controversies*, S. W. Gangestad and J. A. Simpson, Eds. New York, New York, USA: The Guilford Press, 2007.

[36] A. Montagu, "Neonatal and infant immaturity in man," *Journal of the American Medical Association*, vol. 178, no. 1, pp. 56–57, 1961.

[37] K. Nelson, "Evolution and development of human memory systems," in *Origins of the Social Mind: Evolutionary Psychology and Child Development*, B. J. Ellis and D. F. Bjorklund, Eds. London, UK: The Guilford Press, 2005.

[38] J. Lindblom and T. Ziemke, "Social situatedness of natural and artificial intelligence: Vygotsky and beyond," *Adaptive Behavior*, vol. 11, no. 2, pp. 79–96, 2003.

[39] M. Mirolli and D. Parisi, "Towards a Vygotskyan cognitive robotics: The role of language as a cognitive tool," *New Ideas in Psychology*, vol. 29, no. 3, pp. 298–311, 2011.

[40] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.

[41] K. Dautenhahn, "A paradigm shift in artificial intelligence: Why social intelligence matters in the design and development of robots with human-like intelligence," in *50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence*, M. Lungarella, F. Iida, J. Bongard, and R. Pfeifer, Eds. Berlin, Germany: Springer, 2007.

[42] K. Dautenhahn and A. Billard, "Studying robot social cognition within a developmental psychology framework," in *Third European Workshop on Advanced Mobile Robots*, Zürich, Switzerland, 1999.

[43] M. Wheeler, "What matters: Real bodies and virtual worlds," in *SmartData: Privacy Meets Evolutionary Robotics*, I. Harvey, A. Cavoukian, G. Tomko, D. Borrett, H. Kwan, and D. Hatzinakos, Eds. New York, New York, USA: Springer, 2013.

[44] P. R. Smart and K. Sycara, "Situating cognition in the virtual world," in *6th International Conference on Applied Human Factors and Ergonomics*, Las Vegas, Nevada, USA, 2015.

[45] M. Strohmaier and C. Wagner, "Computational social science for the World Wide Web," *IEEE Intelligent Systems*, vol. 29, no. 5, pp. 84–88, 2014.

[46] S. Dieleman, K. W. Willett, and J. Dambre, "Rotation-invariant convolutional neural networks for galaxy morphology prediction," *Monthly Notices of the Royal Astronomical Society*, vol. 450, no. 2, pp. 1441–1459, 2015.

[47] C. J. Lintott and J. Reed, "Human computation in citizen science," in *Handbook of Human Computation*, P. Michelucci, Ed. New York, New York, USA: Springer, 2013.

[48] L. von Ahn and L. Dabbish, "Designing games with a purpose," *Communications of the ACM*, vol. 51, no. 8, pp. 58–67, 2008.

[49] D. Ferrucci *et al.*, "Building Watson: An overview of the DeepQA project," *AI Magazine*, vol. 31, no. 3, pp. 59–79, 2010.

[50] K. Dautenhahn, B. Ogden, and T. Quick, "From embodied to socially embedded agents—implications for interaction-aware robots," *Cognitive Systems Research*, vol. 3, no. 3, pp. 397–428, 2002.

[51] S. K. Chalup, "Incremental learning in biological and machine learning systems," *International Journal of Neural Systems*, vol. 12, no. 6, pp. 447–465, 2002.

[52] K. Sterelny, *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, Massachusetts, USA: MIT Press, 2012.

[53] N. R. Shadbolt and P. R. Smart, "Knowledge elicitation: Methods, tools and techniques," in *Evaluation of Human Work*, 4th ed., J. R. Wilson and S. Sharples, Eds. Boca Raton, Florida, USA: CRC Press, 2015.

[54] C. Allen, W. Wallach, and I. Smit, "Why machine ethics?" *IEEE Intelligent Systems*, vol. 21, no. 4, pp. 12–17, 2006.

[55] M. Walport, "The Internet of Things: Making the most of the Second Digital Revolution," UK Government Office for Science, London, UK, Tech. Rep., 2014.

[56] D. Kirsh and P. Maglio, "On distinguishing epistemic from pragmatic action," *Cognitive Science*, vol. 18, pp. 513–549, 1994.

[57] R. Menary, "Cognitive practices and cognitive character," *Philosophical Explorations*, vol. 15, no. 2, pp. 147–164, 2012.

[58] S.-H. Myaeng, Y. Jeong, and Y. Jung, "Experiential knowledge mining," *Foundations and Trends in Web Science*, vol. 4, no. 1, pp. 1–102, 2012.

[59] N. R. Shadbolt, "Knowledge acquisition and the rise of social machines," *International Journal of Human–Computer Studies*, vol. 71, no. 2, pp. 200–205, 2013.

[60] A. Mao, E. Kamar, and E. Horvitz, "Why stop now? Predicting worker engagement in online crowdsourcing," in *Conference on Human Computation and Crowdsourcing (HCOMP-2013)*, Palm Springs, California, USA, 2013.

[61] A. Segal *et al.*, "Improving productivity in citizen science through controlled intervention," in *24th International World Wide Web Conference*, Florence, Italy, 2015.

[62] A. Segal, Y. Gal, E. Kamar, E. Horvitz, A. Bowyer, and G. Miller, "Intervention strategies for increasing engagement in crowdsourcing: Platform, predictions, and experiments," in *25th International Joint Conference on Artificial Intelligence*, New York, New York, USA, 2016.

[63] B. Settles, "Active learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 6, no. 1, pp. 1–114, 2012.

[64] D. Cohn, "Active learning," in *Encyclopedia of Machine and Data Mining*, C. Sammut and G. I. Webb, Eds. New York, New York, USA: Springer, 2017.

[65] A. Holub, P. Perona, and M. C. Burl, "Entropy-based active learning for object recognition," in *IEEE Online Learning for Classification Workshop*, Anchorage, Alaska, USA, 2008.

[66] L. Barrington, D. Turnbull, and G. Lanckriet, "Game-powered machine learning," *Proceedings of the National Academy of Sciences*, vol. 109, no. 17, pp. 6411–6416, 2012.

[67] F. Ciravegna, S. Chapman, A. Dingli, and Y. Wilks, "Learning to harvest information for the Semantic Web," in *First European Semantic Web Symposium*, Heraklion, Crete, Greece, 2004.

[68] S. Sarawagi, "Information extraction," *Foundations and Trends in Databases*, vol. 1, no. 3, pp. 261–377, 2008.

[69] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[70] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, 2016.

[71] R. Dale, "The return of the chatbots," *Natural Language Engineering*, vol. 22, no. 5, pp. 811–817, 2016.

[72] J. Lester, K. Branting, and B. Mott, "Conversational agents," in *The Practical Handbook of Internet Computing*, M. P. Singh, Ed. Boca Raton, Florida, USA: Chapman & Hall/CRC, 2004.

[73] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with Twitter: What 140 characters reveal about political sentiment," in *Fourth International AAAI Conference on Weblogs and Social Media*, Washington D.C., USA, 2010.

[74] M. Donald, *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, Massachusetts, USA: Harvard University Press, 1991.

[75] A. Clark, "Magic words: How language augments human computation," in *Language and Thought: Interdisciplinary Themes*, P. Carruthers and J. Boucher, Eds. Cambridge, UK: Cambridge University Press, 1998.

[76] ——, "How to qualify for a cognitive upgrade: Executive control, glass ceilings and the limits of simian success," in *The Complex Mind: An Interdisciplinary Approach*, D. McFarland, K. Stenning, and M. McGonigle-Chalmers, Eds. Basingstoke, England, UK: Palgrave Macmillan, 2012.

[77] D. Dennett, *Consciousness Explained*. Boston, Massachusetts, USA: Little, Brown & Company, 1991.

[78] A. Clark, “Material symbols,” *Philosophical Psychology*, vol. 19, no. 3, pp. 291–307, 2006.

[79] ——, “Language, embodiment, and the cognitive niche,” *Trends in Cognitive Sciences*, vol. 10, no. 8, pp. 370–374, 2006.

[80] G. Lupyan, “The centrality of language in human cognition,” *Language Learning*, vol. 66, no. 3, pp. 516–553, 2016.

[81] G. Lupyan and B. Bergen, “How language programs the mind,” *Topics in Cognitive Science*, vol. 8, no. 2, pp. 408–424, 2016.

[82] G. Lupyan, “Carving nature at its joints and carving joints into nature: How labels augment category representations,” in *Modelling Language, Cognition and Action*, A. Cangelosi, G. Bugmann, and R. Borisyuk, Eds. Singapore: World Scientific, 2005.

[83] P. G. Schyns, “A modular neural network model of concept acquisition,” *Cognitive Science*, vol. 15, no. 4, pp. 461–508, 1991.

[84] R. K. Thompson, D. L. Oden, and S. T. Boysen, “Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task,” *Journal of Experimental Psychology: Animal Behavior Processes*, vol. 23, no. 1, pp. 31–43, 1997.

[85] I. M. Pepperberg and S. Carey, “Grey parrot number acquisition: The inference of cardinal value from ordinal position on the numeral list,” *Cognition*, vol. 125, no. 2, pp. 219–232, 2012.

[86] G. Lupyan and E. J. Ward, “Language can boost otherwise unseen objects into visual awareness,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 35, pp. 14 196–14 201, 2013.

[87] A. Clark, “Word, niche and super-niche: How language makes minds matter more,” *THEORIA: An International Journal for Theory, History and Foundations of Science*, vol. 71, no. 3, pp. 255–268, 2005.

[88] K. Friston, “The free-energy principle: A unified brain theory?” *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[89] A. Clark, *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. New York, New York, USA: Oxford University Press, 2016.

[90] ——, “Whatever next? Predictive brains, situated agents, and the future of cognitive science,” *Behavioral and Brain Sciences*, vol. 36, no. 3, pp. 181–253, 2013.

[91] G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[92] K. Friston, “A free energy principle for biological systems,” *Entropy*, vol. 14, no. 11, pp. 2100–2121, 2012.

[93] L. S. Vygotsky, “Consciousness as a problem in the psychology of behavior,” *Soviet Psychology*, vol. 17, no. 4, pp. 3–35, 1925/1979, original work published 1925.

[94] M. S. A. Graziano and S. Kastner, “Human consciousness and its relationship to social neuroscience: A novel hypothesis,” *Cognitive Neuroscience*, vol. 2, no. 2, pp. 98–113, 2011.

[95] N. Phan, D. Dou, H. Wang, D. Kil, and B. Piniewski, “Ontology-based deep learning for human behavior prediction in health social networks,” *Information Sciences*, vol. 384, pp. 298–313, 2017.

[96] C. Vondrick, H. Pirsiavash, and A. Torralba, “Anticipating visual representations from unlabeled video,” in *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, Nevada, USA, 2016.

[97] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, “Traffic flow prediction with big data: A deep learning approach,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.

[98] G. Gómez, M. Lungarella, P. Eggenberger Hotz, K. Matsushita, and R. Pfeifer, “Simulating development in a real robot: On the concurrent increase of sensory, motor, and neural complexity,” in *Fourth International Workshop on Epigenetic Robotics*, Genoa, Italy, 2004.

[99] J. L. Elman, “Learning and development in neural networks: The importance of starting small,” *Cognition*, vol. 48, no. 1, pp. 71–99, 1993.

[100] D. F. Bjorklund, “The role of immaturity in human development,” *Psychological Bulletin*, vol. 122, no. 2, pp. 153–169, 1997.

[101] D. F. Bjorklund and B. L. Green, “The adaptive nature of cognitive immaturity,” *American Psychologist*, vol. 47, no. 1, pp. 46–54, 1992.

[102] M. Lungarella and L. Berthouze, “Adaptivity through physical immaturity,” in *2nd International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Edinburgh, Scotland, 2002.

[103] E. L. Newport, “Maturational constraints on language learning,” *Cognitive Science*, vol. 14, no. 1, pp. 11–28, 1990.

[104] G. Turkewitz and P. A. Kenny, “Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement,” *Developmental Psychobiology*, vol. 15, no. 4, pp. 357–368, 1982.

[105] S. Kirby and J. R. Hurford, “The evolution of incremental learning: Language, development and critical periods,” Department of Linguistics, University of Edinburgh, Edinburgh, UK, Tech. Rep. Occasional Paper EOPL-97-2, 1997.

[106] A. W. Kersten and J. L. Earles, “Less really is more for adults learning a miniature artificial language,” *Journal of Memory and Language*, vol. 44, no. 2, pp. 250–273, 2001.

[107] B. P. Cochran, J. L. McDonald, and S. J. Parault, “Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners,” *Journal of Memory and Language*, vol. 41, no. 1, pp. 30–58, 1999.

[108] Y. Nagai, Y. Kawai, and M. Asada, “Emergence of mirror neuron system: Immature vision leads to self-other correspondence,” in *IEEE International Conference on Development and Learning*, Frankfurt, Germany, 2011.

[109] P.-Y. Oudeyer and L. B. Smith, “How evolution may work through curiosity-driven developmental process,” *Topics in Cognitive Science*, vol. 8, no. 2, pp. 492–502, 2016.

[110] A. G. Barto, “Intrinsic motivation and reinforcement learning,” in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Berlin, Germany: Springer, 2013.

[111] J. Gottlieb, P.-Y. Oudeyer, M. Lopes, and A. Baranes, “Information-seeking, curiosity, and attention: Computational and neural mechanisms,” *Trends in Cognitive Sciences*, vol. 17, no. 11, pp. 585–593, 2013.

[112] P. R. Smart, “Human-extended machine cognition,” *Cognitive Systems Research*, in press.

[113] P. R. Smart, R. Heersmink, and R. W. Clowes, “The cognitive ecology of the Internet,” in *Cognition Beyond the Brain: Computation, Interactivity and Human Artifice*, 2nd ed., S. J. Cowley and F. Vallée-Tourangeau, Eds. Cham, Switzerland: Springer International Publishing, 2017.

[114] P. R. Smart, R. Clowes, and R. Heersmink, “Minds online: The interface between web science, cognitive science and the philosophy of mind,” *Foundations and Trends in Web Science*, vol. 6, no. 1–2, pp. 1–232, 2017.