Foundations of a Philosophy of Collective Intelligence

Harry Halpin [1]

Abstract. Philosophy, artificial intelligence and cognitive science have long been dominated by the presupposition that intelligence is fundamentally individual. Recent work in cognitive science clearly undermines that notion. Increasingly, intelligence is seen not as having its locus in the individual, but in the network of relationships that the individual has with the external world and other individuals. At the same time, there has been an increasing neo-Heideggerian focus on the role of embodiment and anti-representationalism, as shown by work ranging from robotics to dynamical systems. While philosophers are carefully trying to justify this development, the most significant computational phenomenon by far - the World Wide Web - is a veritable explosion of representations. In its latest stage, the Web has become increasingly more the realm of representations used for social real-time co-ordination, as a tool for "collective intelligence." In order to make sense of these developments, we first summarize the differences between the Cartesian assumptions of classical artificial intelligence and the neo-Heideggerian embodied cognitive science. Then we show both how Brian Cantwell Smith's story of representations can be built on top of a neo-Heideggerian story. A combination of a refined version of Smith's rehabilitation of representationalism with the Extended Mind Hypothesis can explain the emergence of collective intelligence and its mediation through representations, and so the wide-scale success of the Web. Finally, we reconsider the notions of autopoiesis, the individual body and embodiment itself in light of collective intelligence.

1 The Individual Challenged

The paradigmatic problem of both analytic philosophy and cognitive science is to explain the intelligence of the human individual: What properties of the individual human deserve credit for intelligence, and why? The answers seem to be self-evident; the unique combination of language and consciousness of the individual is the foundation of intelligence, both of which are not obviously found in ants, trees, or computers. Language and consciousness both seem to be incarnations of a reasoning process that leads to flexible, adaptive behavior, the general purpose reasoning mechanism of Descartes. Ranging from Frege and Russell onwards, philosophy of language sought to explain the relationship of the logical grammar and the world in order to explain why language is so effective, while more recently philosophers have been flocking to the rather mysterious "hard" problem of consciousness. On a more empirical vein, artificial intelligence attempts to understand intelligence through building mechanisms that display intelligence. Yet after the failure of classical artificial intelligence[2] to produce intelligence in computers that could scale out of very small domains, a strain of research based primarily in robotics have shown that the very details of the implementation can produce intelligent behavior without representations, much less consciousness or reasoning [3]. This empirically-driven focus on embodiment has signalled the greatest change in artificial intelligence since its inception, and is explained by Wheeler as a shift from a classical Cartesian paradigm to a neo-Heideggerian programme [32]. Despite this revolution, one assumption that analytic philosophy, classical AI, and the new embodied AI all share is that the fundamental unit of analysis should be the individual.

Recent empirical work in psychology and cognitive science has increasingly challenged the assumption that intelligence is irreducibly individual. It has shown that for complex tasks such as ship navigation that the success of the action relies on the co-ordination of multiple individuals [16]. Evidence from decision-making shows that the "wisdom of crowds" - in other words, decision-making guided by the aggregate of information in a social network - reliably makes better decisions than any individual [4]. Furthermore, work in developmental psychology has shown that the ability to point in children is more than an expression of a linguistic demonstrative, but rather an effort to produce a shared intentionality by directing the attention of others to the same object [31]. Some evidence from neuroscience the explosion of frontal cortex, long thought to be the seat of reasoning, evolved to keep track of interactions within a social network [8], and that the presence of mirror neurons provides a set of neurological mechanisms that allow individuals to share the same neurological state [9]). More recent work in tracking the behavior of individuals finds that their behavior - ranging from movement to turn-taking in conversation - can be reliably tracked by appealing to the behavior of others in their social network with a high degree of accuracy (over 40 to 80% of variation over a wide variety of tasks) without any appeal to planning, reasoning, or verbal language [24]. Pentland claims "that important parts of our personal cognitive processes are caused by the network via unconscious and automatic processes such as signaling and imitation, and that consequently, important parts of our intelligence depend upon network properties." Instead of locating the intelligence in the individual, intelligence can be located in the collective aggregate of individuals.

Collective intelligence does not necessarily mean the sharing of a cognitive state by, for example, mirror neurons. Intelligence can be exhibited by a network of individuals where each individual is specialized in a particular task so that no two individuals share the same cognitive state (skills, activity, and so on) per se, but that the successful action depends on the activities of the entire network. The classic cognitive ethnographic example by Hutchins is the piloting of a ship, where correct piloting of the ship depends on each individual, ranging from the navigator to the steersmen, completing their task [16]. Furthermore, it is not the simple aggregate or organization of individuals in a network that deserves credit for intelligence, but the

[1] University of Edinburgh email: H.Halpin@ed.ac.uk

[2] It should be noted that artificial intelligence is usually the application of reigning theories in philosophy, and classical artificial intelligence was based on the "Language of Thought" representationalism in philosophy of language [5].

conjunction of this social network with their environment. The environment should not be considered static, but dynamically shaped by the actions of intelligent behavior. However, some of the knowledge needed for success is not just embodied in individuals, but embodied in the environment, in their artifacts such as compasses and maps, and the very shape of the boat itself. This leads us to consider the example put forward by Herbert Simon of the apparent complexity of an ant's path as it steadily marches towards food on the beach: "Viewed as a geometric figure, the ant's path is irregular, complex, and hard to describe. But its complexity is really a complexity in the surface of the beach, not the complexity in the ant" [27]. Although this may be true in some cases, it would be to primitive to describe the ant totally to be at the mercy of its environment. Intelligence in general - collective or not - leave traces behind in the environment. The classic example is the pheromone trace of the ant, in which a traces get reinforced as more ants use a particular trail, has been shown to be an efficient way of navigating the environment. This shows how individuals with limited memory can use the shaping of the environment as an external memory. Culture, ranging from design of cities to Wikipedia, can be considered collective cognition extended into the environment. This usage of the environment has a number of advantages over direct individual-to-individual communication. As noted by Heylighen, there is no need for simultaneous presence, so interaction can be asynchronous, and individuals can even be anonymous and unaware of each other. This allows highly organized successful actions to be performed by individual that, due to limited memory and knowledge, would be unable to achieve success otherwise [14].

To modify Pentland's thesis: The collective activity of individuals and their modifications to the environment are responsible for intelligence. While at first this thesis seems intuitive, it goes against much of the practice of both classical cognitive science and philosophy that have a tendency towards individualist reductionism. While the question of whether or not this thesis is actually true is a distinctly empirical question, the philosophical ramifications of this thesis should be developed to see if they are in conflict or continuous with the neo-Heideggerian framework currently being championed in philosophy and AI. Two points of conflict immediately become apparent. Although Heidegger himself is unclear, the neo-Heideggerian framework as articulated by Wheeler understands intelligence as a function of the situated being in the world, not a collective of beings in a shared world [32]. Furthermore, the neo-Heideggerian framework does not explain the reshaping of the environment by intelligence, in particular the creation of representations, not just representational explanation. Representations are seen as crucial by many for the emergence of collective intelligence, which Hutchins traces his "distributed cognition" to "the propagation of representational states across representational media" [16]. The neo-Heideggerian framework is most associated with robotics that exhibits "intelligence without representation," and in contrast collective intelligence is most associated with the advent of the Web, a veritable explosion of representations if ever there was one.

To tackle these problems, we will focus on them in reverse order. First, after explaining the rising neo-Heideggerian framework in cognitive science by contrasting it with the classical Cartesian framework, we will show how representations can be built into the framework. Then, by pushing on the Extended Mind thesis, we will show how the neo-Heideggerian framework allows collective intelligence, including those that use representations. We can then use this framework to understand the explosion of collective intelligence on the representation-heavy Web, and finally try to reconstruct a notion of what should replace the individual in philosophy.

2 Neo-Heideggerian Embodiment

The philosophical assertions made by proponents of neo-Heideggerian programme must be summarized in order to see if they are continuous, or in contradiction with, a theory of representation-based collective intelligence. This is difficult, as like classical artificial intelligence, the move towards embodiment in AI has mainly been one of empirical work where the philosophical assumptions have for the most part been implicit in the work itself. Just as Dreyfus unearthed the philosophical presuppositions of Cartesian classical artificial intelligence, Wheeler has effectively summarized the assertions of embodied AI and based them firmly on a reading of Heidegger, which we call the *neo-Heideggerian programme* [32]. The neo-Heideggerian programme is best understood in contrast with the neo-Cartesian programme of classical AI. Wheeler digests this programme into three main assumptions:

- The subject-object dichotomy is a primary characteristic of the cognizers ordinary epistemic situation
- Mind, cognition, and intelligence are to be explained in terms of representational states and the ways in which such states are manipulated and transformed.
- The bulk of intelligent human action is the outcome of general purpose reasoning processes that work by retrieving just those mental representations that are relevant to the present behavioral context and manipulating and transforming those representations in appropriate ways as to determine what to do

It should be noted that at first glance these neo-Cartesian assumptions are based on the individual being the locus of intelligence. That is surely how at least Descartes thought of it: The singular subject is operative in "cogito ergo sum." The first of the Cartesian points seems to have an implicit individual subject, while the second remains neutral, and the third also seems to have an implicit human individual as the subject. Wheeler then makes the fairly accurate assessment that "word on the cognitive-scientific street is that classical systems have, by and large, failed to capture in anything like a compelling way, specific styles of thinking at which most humans naturally excel" [32]. However, all hope is not lost for AI if it can only lose its neo-Cartesian assumptions. Based on a survey of current work in AI, ranging across robotics, artificial life, and dynamical systems, Wheeler unifies these diverse works on four new assertions, which he states as follows [32]:

- **The primacy of online intelligence**: The primary expression of biological intelligence, even in humans, consists not in doing math or logic, but in the capacity to exhibit...online intelligence...a suite of fluid and flexible real-time adaptive responses to incoming sensory stimuli.
- **Online intelligence is generated through complex causal interactions in an extended brain-body-environment system**: Online intelligent action is grounded not in the activity of neural states and processes alone, but rather in the complex causal interactions involving not only neural factors, but also additional factors located in the non-neural body and the environment.
- **An increased level of biological sensitivity**: Humans and animals are biological systems - and that matters for cognitive science.
- **A dynamical systems perspective**: Cognitive processing is fundamentally a matter of state space evolution in certain kinds of dynamical systems.

Is there any bias towards an individual subject in these assertions? It seems present in a subtle manner in the first assertion since the

very idea of "incoming sensory stimuli" presumes an individual that is processing these stimuli. The second and third assertion also seem to take for granted that our primary subject is not just an individual, but a biological individual. This is put into perspective by the second assertion that "not only neural factors, but also additional factors located in the non-neural body and the environment" play a critical role, a point we will return to with a vengeance.

Wheeler and his philosophical fellow-travellers such as Clark [5] spend much of their time on the question of whether or not there is any room whatsoever for internal representations inside these individuals. Rejecting Clark's notion of "decoupling" as sufficient but not necessary for cases he believes demands a representational explanation, Wheeler argues for some, albeit limited role for representations that pins representations on the two notions of homuncularity and arbitrariness. Since it is too involved to argue over homuncularity and arbitrariness here, we shall instead focus on how Brian Cantwell Smith's revival of decoupability can be built on a neo-Heideggerian framework. We shall just comment that Wheeler's general framework is not incompatible with our notion of collective intelligence and his account of representations is not too far from our account.

3 Representations Revisited

The very idea of representation is often left under-defined and is as a consequence given near-magical powers by certain theories of language and classical AI. While it is hard to pin down a reigning definition, the classic definition stems from the notion of a "symbol" given by Simon and Newell's *Physical Symbol Systems Hypothesis* [22]:

"An entity X designates an entity Y relative to a process P, if, when P takes X as input, its behavior depends on Y."

First, the very idea of "being a representation" is grounded in the behavior of a process, and behavior depends on having access to the representation. Thus, the target of representation (i.e. what is represented, the "thing designated") will depend on the process the representation is used in, i.e. a representation is never context-free. Second, there is clearly decoupling "for this is the symbolic aspect, that having X (the symbol) is tantamount to having Y (the thing designated) for the purposes of process P" [22]. This definition seems to have an obvious point of conflict with the neo-Heideggerian agenda, for it reflects the infamous "subject-object dichotomy" due to its presupposition of at least three distinct a priori entities, the subject (P), the representation (X), and the object (the "target" of the representation, Y). To the extent that these distinctions are held a priori, then the definition is the very exemplar of the neo-Cartesian programme of classical AI.

An escape-hatch from this Cartesian dead-end would exist if there was a way within the neo-Heideggerian program to tell the story of how representations come to be without an a priori subject-object dichotomy. Brian Cantwell Smith tackles this by developing a theory of representations that does not presume an individual [28]. Smith starts with the example from Lettvin and Maturana, a frog tracking a gadfly across the sky [17]. The frog sees the fly, and begins tracking it with its eyes as it flies. The frog and the gadfly are both physically connected via light-rays. Borrowing an analogy from physics, everything is composed of non-distinct fields of energy, so it would be a presupposition to talk about a frog, a fly and light as individual objects. All that exists is some sort of pre-individual flow from which individual objects may emerge. At the moment of tracking, connected as they are by light, the frog, its light cone, and the fly are a system, not distinct individuals. An alien visitor might even think they were a single individual. When the fly goes behind a tree, and the fly emerges from the other side of the tree, the frog's eyes are not focused on the point the fly was at before it went behind the tree, but the point the fly would be at if it continued on the same path.[3] Components of the flux are now physically separated, with a mutually distinct o-region and s-region. The s-region is distinguished from the o-region by virtue of not only its physical disconnection but by the s-region's attempt to "track" the o-region, "a long-distance coupling against all the laws of physics" [28]. After disconnection (and possibly more cycles of disconnection and re-connection) the s-region can stabilize as an individual subject and the o-region as an individual object, and with considerable work on the subject's side to "track" its object a representation is created by the subject using some form of dynamically incoherent memory. Both subject and object are then full-blown individuals, with the subject possessing a representation of the object[28]. The individuals are not a-priori distinct, but co-constitute each other. According to this explanation subject and objects co-evolve, with the physical processes used to track the object being the representation.

In order to clarify and make abstract Smith's analogy and explicitly connect it to Simon and Newell's definition, we can divide Smith's process into what I have called the *representational cycle* [10]. In order to explicate why precisely the s-region differs from the o-region, we rely on Rocha and Hordijk's work on evolving representations, in particular their idea of dynamically incoherent memory [25]. Dynamically incoherent memory is defined as a type of memory not changed by any dynamic process it initiates or encounters. In this manner, it serves as memory that does not degrade or radically alter, but can maintain itself over time. To phrase this outside of the language of dynamical systems, we would say that "dynamically incoherent" might be a misleading word. Instead, what Rocha means is that the subject must have a some sort of memory that is capable of maintaining coherence in terms of its physical structure against, "the vagaries and vicissitudes, the noise and drift, of earthy existence" as Haugeland would say [11]. The cycle can then be put into four stages [10]:

- **Presentation**: Process S is in effective local contact (i.e. physically in contact in space-time) with process O. S is the s-region that evolves into the subject that has the representation and O is the o-region that evolves into the object.
- **Input**: The process S is in local effective contact with coherent memory R. An input procedure of S puts R in correspondence with some portion of process O. This is entirely non-spooky since S and O are in effective local contact. R evolves into the representation.
- **Separation**: Processes O and S change in such a way that the processes are non-local.
- **Output**: Due to some local effect in process S, S uses its local effective contact with R to initiate the local dynamic behavior that depends on R for success.

Smith, and our exegesis of him, has shown it is possible to build a theory of representations based on decoupability and correspondence

[3] While simple physics can do this without any intentionality by making the frog's eyes continue along at the same trajectory, for more complex behavior, such as when the fly is not moving at a constant rate but zig-zagging about, more complex tracking is required. Regardless, the point of Smith's example is that disconnection is required for decoupability and so representation

while not presupposing that intelligent behavior of an individual cognizer depends on internal representations - or that representations - or even an individual - exist a priori at all. Representations are also not everywhere as in traditional representationalism, but instead they are deployed as needed when the relevant behavior requires distal co-ordination. Representations - if not representationalism - is continuous with the neo-Heideggerian agenda. In fact, the very story of representations gives us a way to show how the notion of an individual can emerge from some primordial and undefined Heraclitan flux. Representations are not a Cartesian metaphysical assumption, but arise over time in even a neo-Heideggerian world.

4 From the Extended Mind to the Web

Now that we have shown a plausible story about how representations can be built on a neo-Heideggerian framework, we have to explain how these representations can be used to explain the rise of a robustly representational system like the Web without contradicting the neo-Heideggerian framework. Once this has been done, we can use the current activity on the Web as pointing the way for questioning our conception of the individual, and thereby questioning the bias towards the individual as the fundamental unit of analysis of even the neo-Heideggerian framework. To return to the task at hand, one principle of the neo-Heideggerian agenda put forward by Wheeler is that "online intelligence is generated through complex causal interaction in an extended brain-body-environment system" [32]. We can press on this assertion to make room for the active role for representations in general, and for the Web in particular, "an active externalism, based on the active role of the environment in driving cognitive processes" [6]. Since Smith's representations are not necessarily internal or external to a process, we can remain agnostic as regards whether or not internal representations are necessary or even used by an individual. For example, a representation can be stored in the memory "inside" the head of an agent in some neural state, but it can just as easily be stored outside in a map. The debate over the existence of internal representations is an empirical debate best left to empirical work. However, what is less debatable seems to be the fact that representations at least exist externally from particular agents. After all, finding those representational neural states are difficult, but let us not deny the existence of maps!

In their Extended Mind Hypothesis, Clark and Chalmers introduce us to Otto, a man with an impaired memory who navigates about his life via the use of notes in his notebook [6]. Otto wants to navigate to the Metropolitan Museum of Modern Art in New York City from his house in Brooklyn, but to do so with his impaired memory he needs at least the address. To specify more than Clark and Chalmers, let us say that he needs a map.[4] In order to arrive at the museum, Otto needs a map whose components are in some correspondence with the world he must navigate in order to get to the museum, in other words a representation. Let us say that Otto has in his notebook a map to the Museum of Modern Art that exists for the precise purpose of navigating individuals to the museum. It is hard to deny that a map is representational in the sense we have presented above, as it is a representation whose target is the various streets on the way to the Museum. The map is just an external representation in the environment of Otto, and can drive the cognitive processes of Otto in a similar fashion to the way that classical AI assumed internal representations in Otto's head did. Clark and Chalmers point out that if external factors are driving the process, then they deserve some of the credit: "If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process" [6]. In this regard, the Extended Mind thesis undermines the strict division between internal and external of the agent itself, but again, in a way that is compatible with the neo-Heideggerian framework.

Imagine the world to be inhabited by multiple individuals that can access the same representation. In almost all the original examples that Clark and Chalmers use in the Extended Mind argument, they deploy a single person sitting in front of a computer screen [6]. A more intuitive example would be two people using the Internet to both share a single representation. One could imagine Otto trying to find his way to the Museum of Modern Art, and instead of a notebook having a personal digital assistant with access to a map on the Web. Likewise Inga can have access to the exact same map via her personal digital assistant. Since both Otto and Inga are sharing the exact same representation and because they are both using it in the same manner, Inga and Otto can be said to share at least some of the same cognitive state, due to the fact that their individual cognitive states are causally dependent on accessing the same representation. This representation is the "same" precisely because the digital memory of the computer allows "perfect" copies to an extent as Haugeland explains [11]. However, unlike the lone digital computer, what the Web specializes in is allowing *everybody* to access the same set of representations.

The value of external representations comes with their accessibility, for an external representation that is not accessible when its needed cannot be used to enable online intelligence. It is precisely in order to solve this problem that Tim Berners-Lee proposed a World Wide Web as a universal information space [1]. The primary advantage of the Web is that every representation has a unique name, a URI.[5]. The Web allows each representation to be accessed when needed by using its unique name. Combined with the fact that since the representations are digital and can be communicated in a lossless fashion, the Web allows multiple simultaneous accessing of the exact same representation. Since the Web is a universal space of digital representations, two or more individuals can share the same representation simultaneously. Due to the Extended Mind hypothesis, two or more individuals can then, because of simultaneous access, share some of the same cognitive state.

5 The Web as Collective Intelligence

Much as computation has not remained static, neither has the Web. The Web, as originally conceived by its users, was just a collection of documents connected by hyperlinks, albeit one in a universal information space. These documents were mostly static, being authored and maintained by individuals. Although new pages and links could be added without resort to a centralized registry, the content of the Web was for the vast majority of users was not content that they actually created and added to in any meaningful manner. Within the last few years, a combination of easy-to-use interfaces for creating content and a large number of web-sites that prioritize the social and collaborative creation of content by ordinary users have taken off, leading to the phenomenon known as "Web 2.0," literally the next genera-

[4] In fact, many of us would need a map even without an impaired memory, which points to how widespread this phenomenon is.

[5] Originally the "Universal Resource Identifier," now a Uniform Resource Identifier as given in an updated specification [2] These are exemplified by the familiar format of *http://www.example.org*.

tion of the Web.[6]. This transition from the Web of static hyperlinked web-pages to a more interactive and collaborative medium is more accurately described as a transition from a "Web of Documents" to a "Social Web" [15]. Paradigmatic examples of easy-to-use interfaces would be Google Maps (or even Google Earth),[7] while a paradigmatic example of socially-generated content would be Wikipedia[8]. Furthermore, increasingly these web sites are now being woven into the fabric of the everyday life of more and more people. How many people feel that their intelligence in increased when they have immediate access to a search engine to the Web, a massive encyclopedia available in a few seconds notice?

The Social Web then presents an interesting twist on the Extended Mind Hypothesis extension that we presented earlier. Again, Otto is using a web-page in his mobile phone to find his way to the Museum of Modern Art. While our previous example had Otto using the Web as ordinary Web users did years ago, simply downloading some directions and following them, we now add a twist. Imagine not only that Inga and Otto are using a map-producing Web site that allows users to add annotations and corrections, a sort of wiki of maps. Inga, noticing that the main entrance to the Museum of Modern Art is closed temporarily due to construction and so the entrance has moved over a block, adds this annotation to the map, correcting an error as regards where entrance of the Museum of Modern Art should be. This correction is propagated at speeds very close to real-time back to the central database behind the Web site. Otto is running a few minutes behind Inga, and because this correction to the map is being propagated to his map on his personal digital assistant, Otto can successfully navigate to the new entrance a block away. This (near) real-time updating of the representation was crucial for Otto's success. Given his memory issues, Otto would have otherwise walked right into the closed construction area around the old entrance to the Museum and been rather confused. This active manipulation with updating of an external representation lets Inga and Otto possess some form of dynamically-changing collective cognitive state. Furthermore, they can use their ability to update this shared external representation to influence each other for their greater collective success. In this manner, the external representation is clearly social, and the cognitive credit must be spread across not only multiple people, but the representation they use in common to successfully accomplish their behavior. Clark and Chalmers agree, "What about socially extended cognition? Could my mental states be partly constituted by the states of other thinkers? We see no reason why not, in principle" [6]. How we have extended their story is that socially extended cognition is now mediated by external representations, in particular interactive representations on the Web.

Even this example of brings up points for further consideration. Ordinarily as considered in representationalism as a theory of mind, representations are considered notoriously disconnected from their target, and so while this leaves plenty of room to develop a theory of misrepresentation, it leaves quite a lot of work for philosophers to develop how something like an "internal representation" might have a correspondence with a "target" in the external world. Indeed, this understanding of representationalism as some internal language of thought is precisely what we are *not advocating*, for those philosophical problems among others. What our previous example shows is not that representations are some mysterious language of thought, but as Andy Clark put it, "material symbols" capable of being brought into contact with their equally material targets. While the map may not be the territory, it brings Inga and Otto into contact with the territory.

This leads us back full circle to the Web. For example, the collective editing and use of Wikipedia allows its representations to be increasingly part of of the cognitive system of many people. As representations on the Social Web are updated by increasing numbers of people, each representation is increasingly brought into tighter coupling with both its target and the agent using the representation. As each representation is involved in this process of use and updating is brought into closer and closer cognitive updating with more and more individuals, the representations on the Web are brought into tighter and tighter coupling with what its users formerly considered their individual intelligence, and so leading to the phenomenon widely known as collective intelligence. Indeed, there are now problems as simple as navigating down the street or organizing a social event that many today would have difficulty organizing without access to an interactive mapping Web service or a social networking web site. As users contribute more and more content, the collective content of these web-pages becomes increasingly difficult to track down to individuals. Some of these Web-based tools for collective intelligence have no way to track down the original individual author, others like Wikipedia have sophisticated mechanisms in place to track individual contributions. However, as long as the contribution that the collectively-built web page makes is the sum of more than an individual effort, then the credit must be placed upon the collective content, not the individual author. From the standpoint of the user of the representation, the credit must also not just be placed on the creator of the content, but the very technological infrastructure - ranging from the hardware of high-speed fibre optics and wireless routers to the software of protocol design and web server code - that enables the content of the collectively created web site to be delivered when it is needed. The credit for successfully creating and deploying the cognitive scaffolding is more collective than originally thought! It is also this cognitive scaffolding that provides the ability for distributed individuals to rapidly co-ordinate in near real-time through the modifications of representation, so realizing the definition of collective intelligence given by Levy as "A form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills"[18].

6 Conditions of Collectivity

When one throws even the concept of the a priori individual away, one should seriously reconsider what one is left with. Can we throw away the notion of the individual that just happens to co-incide with what is considered a biological body, the 'common sense' body whose ends happen to coincide with the skin? Obviously upon closer inspection, even the individual biological body is a collectivity, for it is obviously composed of a collective of organs, which are in turn a collective of cells, and so on. If so, should one privilege the biological makeup of certain organs? Evidence from neuroscience in the famous "phantom limb" experiments points to the fact that what our consciousness considers the boundaries of our body does not coincide with our actual biological skin, and that experiments ranging from prosthetic limbs to cochlear implants shows that functionally, non-biological components can be easily considered very much part of the body by the consciousness itself. What we are searching for is then a notion that can define an individual body without resort to making biological tissue some sort of "wonder tissue" as Dennett would put it. One candidate is Maturana's notion of autopoiesis, a more refined notion of the homeostasis that defined earlier cybernetic

[6] A term originally coined by Tim O'Reilly for a conference to describe the next generation of the Web

[7] See *http://maps.google.com* and *http://earth.google.com* respectively.

[8] *http://www.wikipedia.org*

systems [20]. Contrary to the Cartesian assumptions of classical artificial intelligence, in their study of frog vision, what Maturana and others discovered was that the frog's eye "speaks to the brain in a language already highly organized and interpreted instead of transmitting some more or less accurate copy of the distribution of light upon the receptions" [17].

This discovery caused Maturana to reconceptualize the foundations of cognitive science in terms of autopoiesis: that "living organization is a circular organization which secures the production or maintenance of the components that specify it in such a manner that the product of their functioning is the very same organization that produces them" [20]. First, a frog is autopoietic precisely because its internal metabolism is inside a boundary, frog-skin, that defines its organization as a frog. Second, the components, the organs, are inside the frog-skin and self-reproducing. Yet autopoietic systems are not entirely closed, for the frog's consumption of gadflys and other interactions with the environment are done in lieu of maintaining its own organization as a frog, since eating allows it to bring in energy to maintains its metabolism. A frog adapts its interactions to the environment to maintain autopoiesis [23]. The effect of the world upon any autopoietic system is only an effect insofar as it causes the system to adjust itself in order to maintain its own autopoiesis.

The problem with Maturana's notion of autopoiesis is again the very idea of a unity which implies a cell with a membrane or the skin of a frog. The first condition of autopoiesis, namely that the components of an autopoietic system "through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them" suits collective intelligence just fine, as there is no reason a priori why these interactions have to be biological [20]. It is the second part of this definition of autopoiesis that causes us trouble, which is that the components that "constitute it (the machine) as a concrete unity in the space which they (the components) exist by specifying the topological domain of its realization as a network," in other words, "organizational closure" [20]. The problem then is that what autopoiesis explains perfectly is the formation of this topological domain, the crucial fact that cells developed membranes and frogs develop frog-skin that allowed them to separate from their environment. What is lacking is more exploration on precisely how these membranes or boundaries also allow interaction with the outside environment. However, frogs eat gadflies, and grass needs sunlight to grow, and humans use the Web to get directions. How can this bias in favor of the closed system willfully be maintained?

The answer of Maturana and Varela is to introduce the concept of structural coupling to deal with an individual organism's interaction with the environment, a "history of recurrent interactions leading to the structural congruence between two (or more) systems"[21]. Yet instead of two closed systems "perturbing" each other for their mutual autopoiesis, it is easy enough to change perspective to see them as one system maintaining co-evolved autopoiesis. Due to this loophole, even autopoietic systems can become open to the external environment - which after all, are necessary for the system's reproduction - and so open to non-biological organic couplings. The problem with autopoiesis is that precisely as it attempts to get away from a reproductive or genic definition of life that presupposes the individual or the propagation of their genes as the primary feature of life, its definition of organizational closure isolates the system from its environment in a way that prevents the individual from actively assimilating parts of the environment into itself as in the Extended Mind argument. Yet this is incorrect, for closure "has nothing to do with the idea of a materially closed system" since "autonomous systems must be thermodynamically far-from-equilibrium systems, which incessantly exchange matter and energy with their surroundings" [30]. The way out of this dilemma is simple, since closure is used "in its algebraic sense: An operation K exhibits closure in a domain D if every result of its operation yields results within D. Thus, the operation of a system has operational closure if the results of its activity remain within the system itself" [30]. The nub of the problem is that the domain is assumed to be static! Indeed, if the cognitive domain of the autopoietic system can expand to envelop ever more parts that fulfill the two conditions of autopoiesis, then the Extended Mind system can apply to the expansion of autopoietic systems, including heretical bio-social-technological systems.

Despite the biological favoritism of Maturana and Varela, there is nothing inherent in autopoiesis that restricts the components of biology in all possible worlds. The work of Licklider and Engelbart both build from this insight, although they knew nothing of the theory of emergent self-organization, much less autopoiesis as developed by Maturana and Varela. Licklider and Engelbart intuitively grasped that digital computing and representations could easily be part of self-sustaining and intelligent systems. Furthermore, their work led directly to the Internet and the World Wide Web. Instead of aiming to have a machine that is as intelligent as a human individual as in artificial intelligence, Licklider proposed that instead humans and digital computers could couple together closely so that they would become literally symbiotic [19]. Although more work is needed to flesh this case out, it seems there is no inherent contradiction in autopoiesis involving non-biological components. If the individual can be defined via autopoiesis, and to maintain its autopoiesis the individual must increasingly incorporate non-biological components, then the individual is no longer a static, closed system, but an open and dynamic system capable of assimilating and decoupling from various components as it goes in and out of autopoiesis, including digital representations and other biological beings. The obvious objection could be that the biological component is reproducing itself, while the non-biological component is not. Yet is not the reproduction of culture itself reproduction? If so, then humans can be considered not just ways for genes to reproduce, but for our evolving and non-biological technology to reproduce as well.

7 Embodiment Reconsidered

If we now have individuals as open systems that can incorporate non-biological components, do we still have individuals in any useful sense of the term? The main objection to getting rid of the individual would be that the very use of the term embodiment is bound up with that of the biological individual. Before further inspection, the notion of embodiment itself needs to be understood as either simultaneous with or separable from the individual biological body. There does seem something slightly amiss in all the rhetoric of embodiment, as Sheets-Johnstone has pointed out: " the term 'embodied' is a lexical band-aid covering a 350-year-old wound generated and kept suppurating by a schizoid metaphysics" [26]. Everything from a blade of grass to a coffee-cup is embodied in a strictly material sense, and no-one argues otherwise or makes an intellectual programme out of this fact. Embodiment can not just be a synonym for having a physical or material body. What is interesting about embodiment is not the usage of the term as a synonym for the body, but the realization that the context provided by a body can have a causal effect on intelligence. The key word then is "context." N. Katherine Hayles has brought to the forefront that embodiment and the body can actually be spliced into two different concepts. The first, the body, is "always normative

relative to some set of criteria" [12]. In contrast, embodiment is the context that goes along with particular bodies, "enmeshed within the specifics of place, time, physiology and culture...embodiment never coincides exactly with "the body" however that normalized concept is understood. Whereas the body is an idealized form that gestures towards a Platonic reality, embodiment is the specific instantiation generated from the noise of difference" [12]. This is precisely why the notion of embodiment is so difficult for any science to capture, since it is bound up in the very particulars of a given situation that a science or any systematic philosophy must by necessity remove in order to develop any sort of predictive power about future situations and any understanding that applies beyond the here and now. In order to fulfill its role as a science, it is no surprise that cognitive science defined the body by the norm of being bound by the skin. Due to this presupposition, cognitive science has focused more on an a priori "body" than embodiment. Like any fundamentally arbitrary norm, when having to deal with the harsh reality of science, it falls apart. The question returns: If we must construct a body, what kind of body can it be, a body without presuppositions?

As Wheeler relates, when Rod Brooks announced his new paradigm in artificial intelligence based on robotics without representation, in order to distance his positive program for AI from the decades of critique by philosophers, Brooks claimed that at least it wasn't German philosophy [32]. While Wheeler has put together a compelling case that Brooks was in fact doing German philosophy, what we are arguing for is not German philosophy in the vein of Heidegger. To make the case clear, the problem with Heidegger traditionally has not been an emphasis on the biological body. Far from it, since the very Heideggerian notion of the "ready-to-hand" undermines the biological body. Let us look at his paradigmatic example: "the less we just stare at the hammer-thing, and the more we seize hold of it and use it, the more primordial does our relationship to it become, and the more unveiledly is it encountered as that which it is - as equipment. The hammering itself uncovers the specific 'manipulability' of the hammer. The kind of Being which equipment possess - in which it manifests itself in its own right, we call readiness-to-hand" [13]. This readiness-to-hand reveals itself not as abstract knowledge, but as smooth behavior facilitated by the combination of human and hammer. As Wheeler puts it, "the human agent becomes so absorbed in her activity in such a way that she has no self-referential awareness of herself as a subject over and above a world of objects"[32]. At the moment of hammering, given the tight coupling, is it not fair to say that the coupled system of hammer-human is a single system? This is especially true if the hammer is being used in such a way - let's say, to build a house for surviving the cold winter - which is needed for the autopoietic survival of the human agent and his attendant culture, including his hammers in the toolbox. How can even Heidegger himself maintain the biological skin as a crucial boundary?[9] Yet somehow, the very notion of Being is mysteriously tied to the individual human body in Heidegger, and this assumption becomes increasingly uncomfortable, given recent scientific evidence, when refitting cognitive science around on a neo-Heideggerian basis.

To overcome the individual-as-body-in-skin presupposition that is so heavily built into Anglo-American philosophy, what we need is not German philosophy, but French philosophy. French theorists Deleuze and Guattari put forward a concept that can replace the notion of a body: the *assemblage*. In contrast with the individual - even autopoietic - body, Deleuze and Guattari "call an assemblage every constellation of singularities and traits deducted from the flow - selected, organized, stratified - in such a way as to converge artificially and naturally"[7]. Any structural coupling of autopoiesis or instance of the Extended Mind, creates an assemblage. Furthermore, note that this concept is not necessarily disembodied, for the convergence that produces an assemblage can arrive from the "noise of difference," i.e. the context of the world without any abstraction [12]. An assemblage allows us to construct an embodied replacement for the individual body that can keep embodiment while throwing out the individual as an a priori concept. According to Deleuze and Guattari, almost everything in our everyday ontology is an assemblage. In fact, the question then becomes what "bottom-outs" an assemblage, and how to determine if an assemblage exists at a given moment. One further notion brought up by Deleuze and Guattari is that of the *body without organs* that is "under way the moment the body has enough of organs and wants to slough them off, or loses them." The "body-without-organs" allows us to conceptualize bodies as not necessarily biological (i.e. built of organs). More importantly, the body without organs captures the dynamic activity of an assemblage that makes it cast off its previous couplings, and create new ones dynamically in response to its situation. This is the opposite of any statically construed normative body, for the body-without-organs "is not at all the opposite of the organs. The organs are not its enemies. The enemy is the organism" [7]. Let us correct them here: the enemy is not the organism, but the organism as a reified a priori individual.

It seems we have painted ourselves into a corner: If all bodies are collective autopoietic assemblages, then why any assemblages to begin with? If we are throwing away any static body-bounded-by-the-skin, why do the bodies recompose into different collective assemblages, some autopoietic, others not? The answer is in our definition of the body; the source of every conception of the body is inherently normative. Norms do not drop out of the sky as if given to us by the angels; the only scientific story we can tell about norms is evolutionary. As Dennett puts it, all norms must eventually ground out in evolution, although the jury still seems out on whether or not evolution selects genes, individuals, or groups of individuals sharing traits [29]. In a Heideggerian note, the formation of assemblages happens in response to the encountering of problems thrown our way by the world, and our attempt to maintain the autopoiesis of these assemblages as they are faced by these problems, ranging from fleeing sabre-tooth tigers to collectively avoiding extinction of the species. Success in these problems is measured in evolutionary terms, whether or not the assemblage can survive and maintain autopoiesis. As the problems change, so will the assemblages. The assemblage of cells known as the biological human body incorporated the assemblage known as the skin as a solution to problems of heat regulation, evaporation adaption, self-defense, and other problems encountered by cells trying to maintain their autopoiesis. Furthermore, this evolutionary story can be harnessed to explain the emergence of collective intelligence in the forms of the Web. Problems today, ranging from mapping the genome to prevent disease to the co-ordination of production and consumption in a globalized market, are far beyond the knowledge and representations easily accessible without the heavy-duty cognitive scaffolding of the Web. The development of the collective intelligence is the only way to harness the fact that "no one knows everything, everyone knows something, all knowledge resides in humanity" [18]. We can detect the formation of new assemblages, and the representations they utilize and incorporate, by paying attention to the problems that threaten the previously stable assemblages.

[9] Although, we might add that Heidegger does make the human body to be "wonder tissue," by regulating hammers and whatnot to "equipment" and denying Dasein to all but humans. Further explication of this would be illuminating, but is beyond the scope of this paper.

8 Conclusions

Conservatively, what we have argued is two-fold. First, that the notion of representations championed by Brian Cantwell Smith can be built on top of neo-Heideggerian notions of embodiment, and this allows phenomena such as the Web to be brought under consideration and explained as sources of intelligence. Second, and more radically, the assertion that "online intelligence is generated through complex causal interactions in an extended brain-body-environment system" can be pressed in such a way that we can philosophically "come out on the other side" and end up in a world that allows collective intelligence built on top of distributed representations. This allows philosophy to escape from the confines of an overly restricted embodiment that is restricted to the biological body, and so "an increased level of biological sensitivity" in cognitive science should be complemented by an equal sensitivity to the non-biological aspects of intelligence. Humans and animals are systems embedded in a non-biological culture - and that matters for cognitive science. Lastly, while we would not argue against the priority of online intelligence per se, we would hope that it does not miss out the fact that increasingly online intelligence is incorporating the heavy use of representations and other aspects of what has traditionally been thought of as "offline" intelligence. Think about the difference between scavenging for nuts and berries and navigating hyperlinks on the Web to discover a map to the grocery store, for as both Deleuze and McLuhan would note, the Web is the return of the information gathering nomad. This also undermines any methodological insistence on a dynamical system analysis, since dynamical systems have shown trouble in handling anything that appears to be a representation and while they are useful in modelling, they are trapped by their own dependence on initial parameters that may or may not be scientifically illuminating [25].

In a more radical direction, we have questioned the biological body as the useful level of analysis for cognitive science, and so a simplistic version of the neo-Heideggerian embodiment programme as pushed by the work on robotics by Brooks [3]. The body is not given, but is created dynamically as a collective assemblage justified in terms of the problem at hand, where success at the task at hand is grounded out in the normativity of evolution. This has certain resonances with work in continental philosophy, in particular Deleuze and Guattari. Defining intelligence in terms of a fully autonomous agent is not even an accurate portrayal of human intelligence, but a certain conception of the individual human subject, "a certain conception of the that may have applied, at best, to that faction of humanity who had the wealth, power, and leisure to conceptualize themselves as autonomous beings exercising their will through individual agency and choice"[12]. By jettisoning this conception, and maintaining the commitment to a certain necessary degree of embodiment as given by a rehabilitated neo-Heideggerian programme, cognitive science can do justice to complex phenomenon such as the advent of the Web and increasing recognition of collective intelligence. Levy notes that cognitive science "has been limited to human intelligence in general, independent of time, place, or culture, while while intelligence has always been artificial, outfitted with signs and technologies, in the process of becoming, collective"[18]. The vast technological changes humanity has engendered across the world are now reshaping the boundaries of human bodies, and so the cognitive world and the domain of cognitive science. This has been a process that has been ongoing since the dawn of humanity, but only now due to the incredible rate of technological progress, as exemplified by the growth of collective intelligence on the Web, does it become self-evident.

REFERENCES

[1] Tim Berners-Lee. IETF RFC 1630 Universal Resource Identifier (URI), 1994. http://www.ietf.org/rfc/rfc1630.txt.

[2] Tim Berners-Lee, Roy Fielding, and Larry Masinter. IETF RFC 3986 Uniform Resource Identifier (URI): Generic Syntax, January 2005. http://www.ietf.org/rfc/rfc3986.txt.

[3] Rodney Brooks, 'Intelligence without representation', *Artificial Intelligence*, **47**(1-3), 139–159, (January 1991).

[4] K. Chen, L. Fine, and B. Huberman, 'Eliminating public knowledge biases in information-aggregation mechanisms', *Management Science*, **50**, 983–994, (2003).

[5] Andy Clark, *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Cambridge, MA, 1997.

[6] Andy Clark and Dave Chalmers, 'The extended mind', *Analysis*, **58**(1), (1998).

[7] Gilles Deleuze and Felix Guattari, *A Thousand Plateaus*, University of Minnesota Press, Minneapolis, Minnesota, 1987.

[8] R. Dunbar, *The Human Story*, Faber and Faber, London, United Kingdom, 2004.

[9] V. Gallese and A. Goldman, 'Mirror neurons and the simulation theory of mind', *Trends in Cognitive Science*, **12**, 493–501, (1998).

[10] Harry Halpin, 'Representationalism: The hard problem for artificial life', in *Proceedings of Artificial Life X*, (2006).

[11] John Haugeland, 'Analog and analog', in *Mind, Brain, and Function*, Harvester Press, New York, NY, (1981).

[12] N. K. Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics*, University of Chicago Press, Chicago, Illinois, 1999.

[13] Martin Heidegger, *Being and Time*, Blackwwell Publishing, London, United Kingdom, 1927. translated by John Macquarrie and Edward Robinson (1962).

[14] F. Heylighen, M. Heath, and F. Van Overwalle, 'The emergence of distibution cognition: a conceptual framework', in *Proceedings of Collective Intentionality IV*, (2004).

[15] Peter Hoshka, 'CSCW research at GMD-FIT: from basic groupware to the social web', *ACM SIGGROUP Bulletin*, **19**(2), 5–9, (1998).

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

[17] J. Lettvin, H. Maturana, W. McCulloch, and W. Pitts, 'What the frog's eye tells the frog's brian', *Proceedings of the Institute Radio Engineers*, **47**(11), 1940–1951, (1959).

[18] Pierre Levy, *Collective Intelligence: Mankind's Emerging World in Cyberspace*, Plenum Press, New York City, New York, 1994.

[19] J. Licklider, 'Man-computer symbiosis', *IRE Transactions on Human Factors in Electronics*, **1**, 4–11, (1960).

[20] Humberto Maturana and Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, D. Reidel Publishing, Dordecht, 1973.

[21] Humberto Maturana and Francisco Varela, *The Tree of Knowledge*, Shambhala Press, Boston, Massachusetts, 1987.

[22] A. Newell, 'Physical symbol systems', *Cognitive Science*, (4), 135–183, (1980).

[23] Ezequiel Paolo, 'Autopoiesis, adaptivity, teleology, agency', *Phenomenology and the Cognitive Sciences*, **4**(24), 429–452, (2005).

[24] A. Pentland, 'On the collective nature of human intelligence', *Adaptive Behavior*, **15**(2), 189–198, (2007).

[25] Luis Rocha and Wim Hordijk, 'Material representations:from the genetic code to the evolution of cellular automata', *Artificial Life*, (11), 189–214, (2005).

[26] M. Sheets-Johnstone, 'Emotion and movement', in *Reclaiming Cognition*, eds., R. Nunez and W. Freeman, Imprint Academic, (1999).

[27] Herbert A. Simon, *The Sciences of the Artificial*, MIT Press, Cambridge, Massachusetts, first edn., 1969.

[28] Brian Cantwell Smith, *The Origin of Objects*, MIT Press, Cambridge, MA, 1995.

[29] K. Stotz and P. Griffiths, 'Genes: Philosophical analyses put to the test', *History and Philosophy in the Life Sciences*, **26**, 5–28, (2004).

[30] Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, MIT Press, Cambridge, Massachusetts, 2007.

[31] M. Tomasello and H. Rakoczy, 'What makes human cognition unique? from individual to shared to collective individuality', *Mind and Language*, **18**(2), 121–147, (2003).

[32] Mike Wheeler, *Reconstructing the Cognitive World: The Next Step*, MIT Press, Cambridge, Massachusetts, United States, 2005.