Thinking Machines and the Philosophy of Computer Science: Concepts and Principles

Jordi Vallverdú, Editor

Logic and Abstraction as Capabilities of the Mind: Reconceptualizations of Computational Approaches to the Mind

David J. Saab and Uwe V. Riss

Abstract

In this chapter we will investigate the nature of abstraction in detail, its entwinement with logical thinking, and the general role it plays for the mind. We find that non-logical capabilities are not only important for input processing, but also for output processing. Human beings jointly use analytic and embodied capacities for thinking and acting, where analytic thinking mirrors reflection and logic, and where abstraction is the form in which embodied thinking is revealed to us. We will follow the philosophical analyses of Heidegger and Polanyi to elaborate the fundamental difference between abstraction and logics and how they come together in the mind. If computational approaches to mind are to be successful, they must be able to recognize meaningful and salient elements of a context and engage in abstraction. Computational minds must be able to imagine and volitionally blend abstractions as a way of recognizing gestalt contexts. And it must be able to discern the validity of these blendings in ways that, in humans, arise from a *sensus communis*.

Introduction

Behind traditional computational approaches to mind we find the idea that we can simulate the mind as we think we might simulate a chess player by computer programs. This approach is assumes that the human mind is based on a symbolic processing model of cognition. Doing so, we overlook that the way a chess player and a computer approach chess playing are fundamentally different. The human player employs not only sequential logic and his symbolic processing capabilities, but also other capabilities that are described by a connectionist model of cognition. Rather than work through the numerous logical and sequential permutations of possible moves, the human player will recognize larger (i.e., schematic) patterns among the pieces of the chessboard and make his moves based on experience gained over a lifetime of playing. Human players will 'feel' what is the correct move for maintaining an advantage or overcoming a disadvantage, using their intuitive sense derived from schemas based on their long lasting practice.

Even if the machines built in this way that they show comparable results, it does not prove that the human mind and the symbolic machine work in the same way. Indeed human beings (and not only they) possess one fundamental capability that cannot be reduced to symbolic logic manipulation, i.e., abstraction or the capability to develop and employ schemas or recognize gestalt from concrete objects that they find in their environment. These schemas humans evolve through the repeated exposure to similar stimulus as part of our lived experience. Schemas have a duality about them —they are patterns of strongly connected elements of cognition that activate based on salient elements of a particular context and they serve as auto-completion processors, allowing us to perceive a gestalt. These capabilities become apparent in human abstraction. Although abstraction can be analyzed ex-post in terms of logic, e.g., looking for common features, we cannot reduce it to a formal logical process. Abstraction is fundamentally related to schema theory and gestalt theory.

However, the capability of abstraction even goes beyond what we can describe by schemas. We will illustrate this point by way of some examples and explain why it is nevertheless advantageous to work with such metaphorical images. One of our goals is to show the limitations of such images. To this end we will refer to relevance of embodiment and embeddedness and show the relevance of these concepts for the understanding of abstraction. Regarding the latter point we will discuss the works of Heidegger and Polanyi and their philosophical approaches contribute to this understanding. We will follow their analyses to elaborate the fundamental difference between abstraction and logics and how they come together in the mind. The interplay can also be explicated on the basis of paradoxes such as the heap paradox (Keefe, 2000, p. 56) where the approaches of schematic processing and symbolic processing conflict with each other. There are already approaches that rely on gestalt theory, however, they are mainly applied in robotics and not incorporated in the philosophy of mind or computational approaches to the mind.

We will explore how these fundamental processes of abstraction etc. on the one hand and logical inference on the other work together, referring to insights gained from Heidegger and Polanyi such as the distinction of *present-at-hand* and *ready-to-hand* and *focal* and *subsidiary awareness*, respectively. Each of their philosophical approaches facilitates recognition of context in which the salient element of focus is situated. It is the contextualized focal entity that is essential for and evokes meaning within cognition and, hence, understanding in a way that integrates schematic abstract thinking with sequential logic.

The two paradigms even work together in mathematics where we also find an extensive use of abstraction (in the sense that we use here). One example is the abstraction of topological structures, expressed by topological axioms, gained from the analysis of real numbers and other analytic structures. It was Frege (1882) who pointed out that the usage of symbols opens up particularly new ways of analyzing the developing structures, e.g., by gestalt-oriented abstraction. It is this particular capability to abstract from symbolic structure that make up the core of mathematics and not the application of logical rules to axioms and propositions.

Enacting a Rationalist Paradigm for a Computational Mind

In the 1960's AI researchers enacted a research program that attempted to enact the rationalist philosophies of 18th and 19th century philosophers: reasoning was calculating (*pace* Hobbes); mental representations are internal to the mind and indicate a separation from body (*pace* Descartes); all knowledge could be expressed by a universally characteristic set of primitives (*pace* Leibniz); concepts are rules (*pace* Kant); concept rules can be formalized (*pace* Frege); and logical atomism (*pace* Russell) is the means to achieve their goal (Dreyfus, 2007). So certain were AI researchers of their progress and success, that Marvin Minsky, leader of the AI lab at MIT, claimed in 1968: "Within a generation we will have intelligent computers like HAL in the film 2001."[1]

AI researchers were critical of philosophers, who they viewed as lacking sufficient understanding of how the mind works despite centuries of philosophical inquiry. After all, if philosophers did have sufficient understanding, wouldn't they have already devised a practical solution to the problem? Those involved in trying to simulate cognition as part of the AI research agenda failed to consider that they were "hard at work turning rationalist philosophy into a research program" (Dreyfus, 2007). Rationality through logic is a core capability of mind, but it is far from sufficient in explaining either the *significance* or *relevance* of what is identified through logic within a particular context, as pointed out by existentialist philosophers such as Heidegger (1927, p. H. 100).

They ran into the problem of context that was taken up by existential philosophers of the 20th century: How can significance and relevance be represented? Heidegger's famous example of the ontological function of a hammer, excluding its defining relationships to nails and other equipment as well as the skills required to use a hammer in favor of a simplified ontological concept of use for building things, shed light on the complexity of context. For Heidegger, all of these things contributed to the ontology of the hammer. He described the concept of *ready-to-hand* as a way of explaining the complexity of the hammer's *being* and as a way of refuting the 'value characteristics' of an object as its sole definition.

Still, AI researchers, unaware of Heidegger's ontological critique, persisted in their rationalist and reductionist paradigm. If they could only represent a few million facts about objects, the complexity problem might be solved! They had grasped one dimension of context—feature recognition—but failed to see the essential dimensions of *significance* and *relevance* of contextual elements and

---

[1] 1968 MGM Press Release for *2001: A Space Odyssey*.

their associated temporality. When it became clear that their few-million-facts approach was insufficient, they discerned that the problem was a 'frame problem.' For a given context or scenario, there are certain salient elements that take priority over others. If they could describe a scenario (i.e., context) well enough, they could identify essential features that were more relevant and assign them a greater value for computation. Unaware to the AI researchers, Edmund Husserl, who is considered to be the father of phenomenology, had already identified the framing approach (Mingers, 2001). Husserl was also the mentor of Heidegger, who furthered the field of phenomenological inquiry and became one of the most influential philosophers of the 20th century.

Framing, or the use of frames, is both somewhat useful and somewhat problematic. It is useful if one can identify and appropriately value the salient elements of a given context. In fact, the very nature of the frame is intended to do so. However, the frame provides no meta-rules for deciding which frame one should apply to a given context. Which frame is relevant? How does one determine its relevance? The problem of the use of frames for computational approaches to mind is that it sets up a process of infinite regress for "recognizing relevant frames for recognizing relevant facts" (Dreyfus, 2007).

The framing approach lead to an unfruitful path of research, as the boundaries of the framed 'world' became smaller and smaller, ending up as a series of micro-worlds that failed to reflect the real world. After Dreyfus (1972) published *What Computers Still Can't Do*—a heavily Heideggerian critique of the traditional AI of rules, representations, symbols, inferences, and frames--AI researchers began to embrace the Heideggerian problem of embodiment and embeddedness. There were several approaches to the problem, notably Brooks' behaviorist approach, Agre's pragmatist approach, and Freeman's neurodynamic model (Dreyfus, 2007). All three accepted Heidegger's critique of Cartesian internalist representations and embraced the view that cognition is embedded and embodied (Haugeland, 1998).

Rationality and Embodiment

The traditional computational approach to mind or intelligence conforms to a rationalist perspective. It borrows from the analytical tradition in philosophy embedded in the scientific method of the physical sciences that tends towards reductionism. AI researchers of the 1960's adopted a view of cognition as a physical symbol system in which the neural activations and impulses of our brains became the analog for computation of symbolic bits of data (Newell & Simon, 1988). These computational approaches assume a one-to-one correspondence of concrete objects in our environment and their symbolic representations so that the comprehension of the world, i.e., intelligence, allows us to reduce the workings of the mind to a simple formula:

Input (presented as abstract symbols) + Algorithm (logical inferences) = Output.

This formula is similar to the way in which AI researchers had thought we processed and used language. We have some words (input) and we apply some forms of inferences and rules to the pattern of lexical units (algorithms) from which we create meaning (more algorithms?) and translate that meaning into another set of lexical units (more algorithms?) that take the form of an expression (output) which also serves as the input for someone else's language processing. It is a very reductionist approach, which found its expression in the Communication Theory of Shannon (Shannon, 1948; Shannon & Weaver, 1949).

The simple formulation of *Input + Algorithm = Output* is more complex even than it may originally seem when we apply it to a purely symbolic system such as language. Meaningful communication depends not only on lexicon and syntax, but also semantics and context. It might be appropriate to apply a mathematical reductionist approach to the signal processing of various mechanical sensors, but its application to language is inadequate. Language may be a symbolic system, but the meaningfulness of those symbols depends upon the experiential understanding of the communication participants. Language is part of our embodied and embedded experiences, not simply strings of lexical units that exist in a semiotic relationship with concepts.

What does it mean to be embedded and embodied? To be embedded means that we, as humans, are not separated from the world. In fact we are inseparable from the world. Wherever and whenever we are, we are embedded in a world, a physical universe, from which our physical being can never escape. We cannot experience anything without the world in which we are embedded. Which brings us to the issue of embodiment. Descartes posited a separation of mind from body, a notion that has had powerful influence on Western philosophy and Western thinking. The Heideggerian critique rejects this separation—'we' are not separate from or separable from our bodies. 'We' are not subjective entities inhabiting an objective body. There is no separation between our subjective mind and our objective bodies. We are embodied, just as we are embedded, and can never escape our embodiment. Perhaps Merleau-Ponty said it most succinctly:

> *In so far as, when I reflect on the essence of subjectivity, I find it bound up with that of the body and that of the world, this is because my existence as subjectivity is merely one with my existence as a body and with the existence of the world, and because the subject that I am, when taken concretely, is inseparable from this body and this world. (Merleau-Ponty, 1962)*

These notions of embeddedness and embodiment have important implications for our cognitive models, which have heretofore been based on the idea that our mind holds 'internal representations' of the external world. The internalist notion of representation propelled early AI researchers to reproduce these representations symbolically within machines. The richness of our environments made this representational approach extremely difficult. With the notions of embeddedness and embodiment, however, we came to the realization that humans avoid the problem of internalist representation "because their model of the world is the world itself" (Dreyfus, 2007).

Brooks was one of the researchers who adopted this non-internalist stance (Brooks, 1988) by constructing robots that act with 'swarm intelligence' that emulated insect-level interaction. The problem, as Dreyfus (2007) points out is that Brooks' robots "respond only to fixed isolable features of the environment, not to changing context or significance." From a Heideggerian perspective, *coping*—our way of dealing with the world in which we are embedded via our embodiment—is more than understanding in terms of inferential symbolic representations and more than Skinnerian responses to fixed features of our environment. Our coping entails an *openness* of our being to the world, which allows us to respond to salient features of our environments without having to attend to the non-salient, but which doesn't exclude our ability to organize our world semantically or express our understanding of it linguistically (Heidegger, 1927, p. H. 163).

Semantic organization includes both the capacity for logical organization and structuring as well as abstraction. Before we continue with our discussion of embodiment, we turn to *abstraction* and show in more detail how abstraction can be understood against the described background.

Abstraction: Transformations from concrete to abstract and back again.

We start with the traditional idea of abstraction and the objects to be abstracted. Abstraction was understood as a psychological process that associates objects, which are part of our experience, with concepts. This association is achieved by attention to those features of the respective concrete object which are common to all objects associated with the respective concept (Honderich, 1995). This notion of abstraction suggests that an abstract object x is defined by a finite set of features $F_x = \{ f_{x,1}, \ldots, f_{x,n}\}$. In the same assumption a concrete object c is regarded as a (quasi) infinite set of features $F_c = \{ f_{c,1}, f_{c,2}, \ldots\}$. Consequently this suggests that an abstraction can be understood as subset formation, i.e., the concrete object c is associated to the abstract object x if $F_x \subset F_c$. According to this scheme we can also define generalizations of abstract objects, i.e., an abstract object g is a generalization of an abstract object c if $F_g \subset F_c$. In this way abstraction becomes a process of logical reasoning that can be performed in a sequential manner.

Although this approach is mathematically very convenient the aforementioned counterarguments suggest that it does not describe abstraction in a proper way. In fact, there are several arguments that raise doubts about its suitability. One argument against the described feature notion of abstraction comes from Wittgenstein (1953) and is known under the term *family resemblance*. It says that the concrete objects associated with an abstracted object are united by a *network of overlapping but discontinuous similarities* (Honderich, 1995) and not by a unique set of features. Another argument originates from the observation of concrete abstraction processes. For example, if we ask a person, who stands in front of a table, what the object in front of her is, she might immediately answer that it is a table. If you ask the same person why it is a table she will probably answer that it is obvious. If you insist on an explanation she will perhaps answer that it has a board and four legs. Obviously this is not a collection of necessary features since people are aware of the fact that there exist tables with fewer legs. Such observations do not support the idea that the observer strictly checks features before she associates a concrete object with an abstract concept. Moreover we can exclude the feature approach because of its tendency towards infinite regress—if the abstraction of a table requires the abstraction of the legs and board, the person would have to identify the respective abstraction features for legs and board as well, leading to an infinite regress. A third argument refers to the continuity of processes. Let us assume a thought experiment in which we transform an object, e.g., a table, continuously into another object, e.g., chair. For example, we could do so by removing tiny pieces from one place and add them to another. In this way we obtain a continuous transformation of the table into the chair and vice versa. Tables and chairs, however, should be characterized by different sets of features. This means that there should be specific points in the transformation process at which the respective features as 'switched on' or 'off'. Since the process is continuous this does not seem to be reasonable. This latter argument is closely related to the heap paradox (Keefe, 2000).

Psychology shows that the association with a certain concept depends on the situation you are starting from (Fisher, 1967). Returning to the aforementioned example of the transformation of a table into a chair, observers recognize the appearance of the table or chair at different stages of transformation depending on what object they start with. When that switch happens is dependent upon the observers' starting points, i.e., whether the transformation begins with the table or the chair. The transformation example indicates that abstraction is a dynamic and context dependent process and not a static mapping of concepts and features. The philosophical conceptualization of abstraction has to take this dynamicism and context dependency into account. This dynamicist perspective encompasses the idea that concepts are learned and that learning is constrained by a capacity to subsume concrete objects within the boundaries of an appropriate concept. The concept of capacity that we refer to here is the same articulated by Aristotle: such capacities become manifest in acts that actualize them (Kern, 2006). For example, a person can actualize her capacity to associate a concrete object with a concept by subsuming such an object under this concept and drawing successful conclusions from the abstract object. This means that the actualization is closely related to concrete acts of the person who possesses this capacity.

If process of learning a new concept starts with an abstract definition, e.g., a table is a board with four legs upon which things are placed. The capacity for this type of abstraction relies on other capacities, i.e., to identify table legs and boards and what it means to deposit things. Over time and through repeated actualization of the association of concrete object with concept, the clustering of associations becomes independent of the respective underlying sub-capacities.

The respective composite abstract turns into simple abstract, i.e., the abstract that was initially defined by multiple features becomes a gestalt. Thus, the person possessing the respective capacity is enabled to identify concrete objects as tables even if they have fewer than four legs. The abstraction, which has started with a feature-based definition, has turned into an embodied capacity. Other concepts, which are not frequently actualized, remain dependent on underlying concepts. So, in other words, if we need the features of the abstract to define or recognize it, they will remain as part of the composite abstract, but if we don't need them, the simple abstract becomes more useful and more frequently used.

Intentionality: An object is an object, but isn't an object, per se.

We recognize that non-logical capabilities are essential for input processing (e.g., perception) as well as output processing. In the act of 'thinking' human beings use both abstraction and logics, which appear as analytic and embodied capacities. On the other hand they perform actions according to 'instructions' that are rooted in embodied capacities. These 'instructions' are logical in a strict symbolic sense, but rather embodied capacities that are developed through practice and repeated exposure to the same or similar stimulus. Elite athletes exhibit such embodied capacities as part of their practice. They often first study how to move optimally before they then train for the intended bodily movements by repeated practice. Through this repeated practice they habituate their neural networks to carry out those movements without needing to engage their capacity for logic and rational thinking. When engaged in their athletic activities, they are able to combine their logical and embodied capacities to act (i.e., intentions) strategically and responsively within a competition environment.

In coping with our everyday existence we encounter two basic modes of intentionality according to Heidegger: (1) an objective intentionality corresponding to the *present-at-hand*, and (2) a deictic intentionality responding to the *ready-to-hand* (Agre, 1988). That which is *ready-to-hand* is more appropriately characterized as the holistic affordability for action that surrounds an object rather than discrete characteristics or qualities of an object. What is *ready-to-hand* constrains the temporal paths of possible action one might take based on the salient elements of an object or situation that has become the focus of one's attention as the *presence-at-hand*. We must not mistake, however, the *present-at-hand* or the *ready-to-hand* for objectivity. Our embeddedness and embodiment preclude a state of objectivity, per se.

Computational approaches to mind have difficulty in not-objectifying that which is *ready-to-hand*. What is *ready-to-hand* is by definition context-dependent, but also involves the possible responses to what is *present-at-hand*. Whatever responses or possible actions are afforded within a context require flexibility, simply because no context or situation is ever the same by virtue of its temporality. There is always something different, even if it is only the semantic organization of experience or learning 'within' *Dasein*[2] that has occurred in the interim.

The fundamental dynamic nature of an embedded and embodied coping is described as *coupling* (van Gelder, 1997; Winograd & Flores, 1987). What we normally think of as cognition, flowing as it does from a symbolic processing model, belies the nature of 'the couple'—of coper and the world. Rationalist approaches favor the symbolic processing approach, while the existentialist approaches view the symbolic processing capabilities of cognition as emerging from the 'dynamical substrate' of coupling:

> *Cognition can, in sophisticated cases, [such as breakdowns, problem solving, and abstract thought] involve representation and sequential processing; but such phenomena are best understood as emerging from a dynamical substrate, rather than as constituting the basic level of cognitive performance.* (van Gelder, 1997, pp. 439, 448)

The representational dimensions of symbolic processing "presupposes background coping [and] depends upon a background of holistic, nonrepresentational coping [that] is exactly the Heideggerian project" (Dreyfus, 2007).

The Enigmatic Nature of Schemas

In cognitive science, connectionist theory posits the human conceptual system as a network composed of a large number of 'units' joined together in a pattern of connections (Rumelhart & McClelland, 1986). Cognitive anthropologists and educational psychologists refer to these patterns of connections as *schemas* (Anderson, Spiro, & Montague, 1984; D'Andrade, 1995; Davis, 1991; Strauss & Quinn, 1997). Schemas are strongly connected networks of cognitive elements, having a bias in activation through repeated exposure to the same or similar stimulus, but they are not rigid and inflexible. They are adaptable, sometimes resulting in the strengthening of existing schemas, sometimes in their weakening in the face of new experience. D'Andrade (1995) explains in more detail that schemas are "flexible configurations, mirroring the regularities of experience, providing automatic completion of missing components, automatically generalizing from the past, but also continually in modification, continually adapting to reflect the current state of affairs." Describing them as 'flexible, mirrored configurations' implies that schemas are structural entities within cognition that are comprised of several elements. Schemas are not the individual elements rather strongly connected clusters of elements of experience within cognition. Elements of experience are clustered in cognition, in our neural networks, because they are clustered in our lived experiences. Clustering cognitive elements makes them more efficient by reducing the cognitive load associated with processing experience.

---

[2] Heidegger coined *Dasein* (literally "there-being") as a way to describe man's way of being in the world. *Dasein's* openness to the experience of being is characterized by *understanding*.

Schemas are powerful processors of experience, help with pattern completion, and promote cognitive efficiency. They serve to both inform and constrain our understanding of experience. People recall schematically embedded information more quickly and more accurately (DiMaggio, 1997). In fact, schemas hold such sway in our cognition that people may falsely recall schematically embedded events that did not occur. They are more likely to recognize information embedded in existing schemas because of repeated activation of the schemas. This repeated activation evokes expectations within cognition and the easy recognition of contradictory or challenging information that does not conform to those expectations formed as part of the existing schemas. Information that is orthogonal to existing schematic structures, that doesn't acquire salience through the repeated activation of schemas and the creation of associated expectations, is much less likely to be noticed or recalled. Because of their functionality in pattern completion, schemas function, in some sense, as flexible filters of experience, enabling us to attend to its salient features while filtering out the non-salient. Schemas allow us to perceive gestalts and help us to limit informational overload.

Schemas don't exist in isolation as objective patterns of neural activation such that they can be plotted on a representational map of a neural network. Schemas are emergent entities that are undergoing subtle changes within a complex network of neural activations that span cortical, limbic and peripheral neural pathways and contribute to our sense of embodiment. Rather than passive receivers of bits of information, our embodiment and embeddedness require us to actively engage the world, to create a *lived experience* (pace Heidegger). The motivational force of some schemas that are activated by salient elements of our environment is what Freeman (1991) refers to as an *attractor landscape*. It is not the particular elements that activate our schemas, but rather their salience—"the significance of the stimulus," (Freeman, 1995). Freeman's research in neurophysiology leads him to the following conclusion:

> *I conclude that context dependence is an essential property of the cerebral memory system, in which each new experience must change all of the existing store by some small amount, in order that a new entry be incorporated and fully deployed in the existing body of experience. This property contrasts with memory stores in computers…in which each item is positioned by an address or a branch of a search tree. There, each item has a compartment, and new items don't change the old ones. Our data indicate that in brains the store has no boundaries or compartments. …Each new state transition…initiates the construction of a local pattern that impinges on and modifies the whole intentional structure. (Freeman, 1995)*

Freeman wants us to imagine a conceptual landscape as if it were a physical landscape with craters. These craters represent concepts, with salient, permeable boundaries that form the rim of the crater. The crater is what Freeman refers to as an *attractor*. And the basin (lowest point) of the crater is a *basin attractor*, which is the conceptual place that it takes minimal energy for our attention to flow.

Now imagine that these craters exist in relation to one another, forming a complex network of basins in the landscape, i.e., an *attractor landscape*. When we view the attractor landscape, we see a vast network of basins, clusters of basins, basins within basins, and basins overlapping basins. Moreover, this landscape of basins lies upon a malleable surface that allows for changes in the landscape based on newly lived experience. Because the entire complex network landscape of craters is interlinked, localized changes arising from experience will have an effect on the structure and strength of the entire network. The attractor landscape metaphor reflects the notion that concepts (i.e., craters) don't exist in isolation but rather as part of the network of schemas we develop through our lived and embodied experiences.

What Freeman postulates is that new conceptual stimuli will impact the attractor landscape and modify its whole structure. Sometimes these conceptual changes obliterate previous topological relationships, resulting in a wholly new localized intentional structure. Sometimes these conceptual changes are more incremental, resulting in a richer local topology of multiple basins within a crater. What's important is that it is not the stimulus, per se, that modifies the attractor landscape, but rather the class and significance of the stimulus for the subject and its effects on localized networks of craters and the relative depths of their basins brought on by new experience.

> *Freeman's model instantiates the causal basis of a genuine intentional arc in which there are no linear casual connections between world and brain nor a fixed library of representations, but where, each time a new significance is encountered, the whole perceptual world of the animal changes so that the significance that is directly displayed in the world of the animal is continually enriched. (Dreyfus, 2007)*

Schemas and the dynamical substrate

Generally schemas are networks of cognitive elements, which we correlate here with abstracts for the sake of simplicity. Henceforth we will use the term schema to describe a flexible representational structure that allows for contextual varieties. What this means is to be explained in the following. Any abstracts we can identify never appear independently. If we imagine a cherry we usually think of a red cherry. Such associations reflect relations between different abstracts that are important for the way we perceive the world. The resulting network of an abstract with other abstracts is considered as a schema. If we consider these relationships we find significant differences in their strength. For example, the relationship between 'cherry' and 'red' is stronger than the relationship between 'cheery' and 'green' while the latter is again stronger than that between 'cherry' and 'purple'.

While schemas represent the wholeness of such relationships, the actual salient representation defined by relevant relationships depends on the particular context in which this schema is expressed. While a schema can include a multitude of possibilities, in a concrete situation only a limited number of relations are relevant (Hagengruber & Riss, 2007). In Heideggerian parlance, this limited set of possible holistic affordances that finds its expression is some of the relevant relations becomes that which is *ready-to-hand*, as we discussed earlier. For example, in the context of art a 'purple apple' might possess some relevance whereas in the context of a supermarket the same relationship would raise astonishment and suspicion whether the respective object is a real apple. Our schemas constrain the *ready-to-hand* possibilities of an apple's color through expectations that become integral to our schemas as part of our patterned experiences.

On the one hand, the schema is compatible with the usual idea of representation of abstract objects by their properties and possible relations to other abstract objects. On the other hand, it is also compatible with Wittgenstein's idea of family resemblance since it allows the same abstract to appear in different relationships to other abstracts depending on the specific context. Relationships are not simply switched on and off but the 'strength' of a relationship can continuously change if the description or experience of the context continuously changes.

The concept of schema helps us to better understand abstraction and the role features actually play in it, i.e., why features can be used to represent abstracts at all and they also explain why a static and context-independent schema is insufficient in explaining abstraction. We start with a fixed model (table = board + 4 legs), and only later through concrete experience realize that it can vary (3 legs). At that point, we draw upon some relevant elements of the schema and blend them with new elements of experience. The blending can vary in significant ways—string from ceiling, single leg, three legs, etc. The schema in this instance is representative of the relationships of features. On the other hand, schemas can be considered to have inherent variability such that they can be representational structures that express themselves differently depending on the context. The context dependency of a schema represents the independence of the considered abstract (i.e., the table) of the related abstracts (i.e., the legs etc.). In a schema, where the relationships are completely context independent, reduction becomes possible. For example, if we state that 'bachelor' is always identified with 'single male', the former concept can be defined (and replaced) by the latter. In contrast, for other concepts such as 'game' as Wittgenstein has shown such a definition is not generally possible. Thus, the more variable a schema becomes the more difficult it is to introduce analytic definitions.

However, schemas are only auxiliary constructs that allow us to illustrate principle processes in the mind. Due to their context dependency and variability they are not suitable for usage as knowledge representation as we find them in finite relational structures of abstracts and as we use them in analytic mental processing. Conversely, when we introduce new abstracts we start with a finite definition, e.g., that a 'table' is a 'flat board with four legs', in which the table is defined as a rigid and context independent schema. In the course of time other *qualia* of tables and their context dependency comes into play and the schema becomes variable. In this process the analytically defined abstract 'table' becomes a concept that is only loosely related to the original definition. In the next section we describe parallels of this process that we find in Polanyi's philosophy.

The irreducibility of tacit knowledge

Taking such embodied capacities into account, Polanyi (1962) has claimed that human knowledge is mainly tacit. This means that this knowledge cannot be verbally expressed in a way that makes it communicable nor can it be reduced to logical processing. In particular this means that tacit knowledge cannot be formalized or represented in a symbolic way. Consequently tacit knowledge cannot be learned by communication but must be acquired by practice. Tacit knowledge describes a human's particular capacity to perform a specific activity, which is learned by actualizing the capacity as part of an activity. We claim that abstraction is mainly based on tacit knowledge and therefore is not logically specifiable (Polanyi, 1962, p. 56). For example we can recognize a face although we cannot describe which parts of the face determined this recognition. We abstract objects in a way in which we perceive the whole without full awareness of the individual features.

To describe this phenomenon Polanyi introduced the distinction between subsidiary and focal awareness that he explained on the basis of the example of a hammer (Polanyi, 1962). If we use a hammer to drive a nail into the wall our primary or focal awareness is connected to the actual process. Nevertheless we are also aware of the hammer, the hand, the nail, etc. but in another, more hidden or subsidiary way. Every time we focus on a whole, our awareness of the parts is subsidiary. If we turn our focal awareness to the hammer or the nail, the awareness of the process becomes secondary and the execution becomes clumsy and less experienced. This is very close to the idea of gestalt, of which we also lose sight if we concentrate on one of its details. However, these details point to a network of related concepts that can be analyzed but also to aspects 'beyond' the network. For example, the shift of focal awareness, whether volitional or non-volitional, describes a phenomenon that cannot be explained by schemas.

The focus of our awareness is related to the activities in which we are currently involved. For example, if we are busy driving a nail into the wall, we concentrate on this particular activity. If we want to analyze the respective sub-activities, our focus moves from one involved object, hand, hammer, or nail, to the next. We realize that it is our intellectual interest that determines the focus but not our bodily involvement in activities, i.e., in interactions with the world. In this way we are describing an ex-post analysis of the activity (e.g., 'hammering'). Activity could also encompass everything, including 'thinking.' On the one hand, we have different foci, which are opposed to activity. On the other hand, the activity could be considered a focus. We can decompose an abstract into different elements and can apply this decomposition to activities, where hammering can be a gestalt that can also be deconstructed into several activities. Abstraction in this example is not fundamentally different from what we consider to be primarily bodily activities.

The example of how a single abstract can emerge from the relational structure of several abstracts through repeated actualization, as we described above, reflects how we learn actions. If we use a hammer for the first time to drive a nail into the wall, we will first concentrate on how to deal with the weight of the hammer; we will experience how a blow of the hammer affects the nail, and so on. All these elements are not yet connected. Over time after some practice, however, these individual elements are merged to one action of driving the nail into the wall.

This description also depicts how we perform abstractions. When we see a table we are subsidiarily aware of its features but not in a way that we can directly name them. Moreover, the capacity that is actualized in this abstraction does not depend on a fixed set of features but rather resembles intuition than logical inference. The variable collection of features, which appear in our subsidiary awareness and can be determined ex-post, is expressed in schemas. Here it must be remarked that the features that we call from subsidiary to focal awareness do not form a fixed set but rely on further abstractions. When we try to find out why the object in front of us is a table we must do two things: first, we have to identify those objects that contribute to the tables, and second, we must abstract these parts again.

Polanyi's approach also bears consequences for communication. Thus, Walsham (2005) pointed out that human communication requires sense-giving by the sender of the message, as a process of abstraction, and sense-reading by the recipient. Both processes are based on a shared understanding of what is meant which again is based on embodied capacities. For example, a person can read Einstein's formula $E = mc^2$ but if she does not possess any experience in physics it will remain a meaningless expression. If the respective person possesses some basic knowledge of physics she may at least know the concepts of energy, mass and velocity of light. In this case the person will be able to bring these concepts together in an abstract way, represented by the abstract formula, but does not understand its practical consequences.

Sensus communis

As beings in the world, we organize our experiences in ways that ensure ease of interaction, coordination of activities, and collaborative interaction. Because we organize our experiences in particular ways, people in the same social environment will indeed experience many of the same typical patterns. In experiencing the same general patterns, people will come to share the same common understandings and exhibit similar emotional and motivational responses and behaviors. However, because we are also individuals, there can be differences in the feelings and motivations evoked by the schemas we hold. "The learner's emotions and consequent motivations can affect how strongly the features of those events become associated in memory" (Strauss & Quinn, 1997, p. 6). Individuals will engage the external world structures and experience the same general patterns. Similar stimuli and experiences will activate similar schemas. It is in that sense we considered them shared schemas. The sharing of schemas does not require people to have the same experiences at the exact same time and place, rather that they experience the same general patterns. It's their quality of sharedness that makes them a dimension of the cultural and from which we derive our *sensus communis.*[3]

Shared or cultural schemas have other qualities also. Some schemas are durable. Repeated exposure to patterns of behavior strengthens the networks of connections among the cognitive elements. Some schemas show historical durability. They are passed along from one generation to the next. Some schemas show applicability across contexts. We draw upon them to help us make sense of new and unfamiliar experiences. Some schemas exhibit motivational force. Such motivation is imparted through learning, explicitly and implicitly, strengthening the emotional connections among the cognitive elements.

We share the intrapersonal dimensions of culture when we interact with others. In sharing these intrapersonal dimensions, schemas are activated. Activation evokes meanings, interpretations, thoughts, and feelings. We make meaning of our experience. The cultural meaning of a thing, which is distinct from the personal cognitive meaning, is the typical interpretation evoked through life experience, with the acknowledgement that a different interpretation could be evoked in people with different characteristic life experiences. In some cases our experience is intracultural, where we share a similar cultural frame. In other cases our experience is intercultural, where we are sharing different cultural frames. The meanings evoked by one person in relation to a particular extrapersonal structure may not be the same as those evoked in another. In fact, the meanings evoked may not be the same within the same person at different times, for they may experience schema-altering encounters in the interim.

Knowledge representation and its limits

Finally we have to answer the question how we actually come to the idea of abstraction as a feature-based analytical process. This answer is that in reality tables are usually well enough distinguished from chairs, cupboards, etc. Although we can construct transitions between these objects in thought experiments, the transition is insignificant for practical purposes. The few exceptions that we find can be handled in an explicit way. However, this insight does not allow us to reduce abstraction to schema evaluation. It only helps us to rationalize abstractions *ex post* if this is necessary, e.g., if a contradiction has

[3] *Sensus Communis* is meant here in the Gadamerian sense—the whole set of unstated assumptions, prejudices, and values that are taken for granted; the non-reflective judgments and values learned but not judged.

occurred. The contextuality and consequent variability of schemas makes it practically impossible to use them for explicit knowledge representation.

Moreover, there are some natural objects that are not so clearly distinguished, e.g., colors or artworks. Here we find abstracts that depend on the particular society. In particular we find that societies that live in different environment tend to different abstractions. This means they use those abstracts that are most likely to be helpful in this environment. These abstractions cannot be learned solely by direct experience but rather by communication and are based on the *sensus communis*. Since the environment is changing, often by human interference, continuously experienced transitions that did not appear originally come into being so that the concepts become less disjoint. In order to deal with such situations humans have to make a cognitive shift from subsidiary awareness of the respective features to focal awareness and analysis of them. This, of course, requires the insight that the respective abstract no longer adequately describes the altering concrete object. For example, the decrease of a heap by removing grains finally leads to the inadequacy of the concept of heap while the respective process can only be understood by considering the number of grains (as a subsidiary feature).

According to Polanyi the fact that we are actually able to do such analysis of abstracts by transition from subsidiary to focal awareness, reflects an ontological structure, which he calls ontological stratification (Polanyi, 1969). This is not to be understood as synonymous with reducibility. Polanyi uses the example of physics and chemistry to explain the relation. Although chemistry cannot be reduced to physics it is nevertheless possible to explain certain chemical transitions, e.g., chemical reactions, by means of physical consideration.

Future Trends

Obviously the traditional proceeding of computational approaches to the mind has to be replaced by other approaches that take the variability of schemas and the embodied embeddedness of the human mind into consideration. The process of abstraction is a perfect example for this requirement. As we have seen fixed schemas can serve as starting point for the formation of a concept but they are not sufficient to deploy the full power of human concepts that are highly adaptive to different concrete situations. The transition from fixed to variable representations results from interaction of humans with their environment. This has to lead to the replacement of static representations by dynamic schemas, which provide capacities that enable the machine to abstract objects that deviate from standard forms.

Moreover even the variability of schemas comes to a limit if we aim at a complete simulation of the human mind. Here we have to take into account that the evolution of abstraction-related capacities is based on structures that even go beyond schemas and reach layers that include more basal bodily systems, such as emotions. Here we hit upon a fundamental problem that consists in the incoherency of bivalent abstracts and continuous processes as they become apparent in the heap paradox. The bivalence of abstracts refers to the fact that we have to associate a concrete object either with an abstract *c* or its negation *not-c* in order to apply logics (otherwise we conflict with the law of the excluded middle) and traditional computational algorithms.

Such an approach requires a certain openness of the machine to new experiences that are not covered by given definitions. In human societies such experience is passed from one individual to another and backed by a *sensus communis*. This means that human beings do not acquires all these capacities on their own but that they learn most of them through communication. This does not mean that knowledge is simply copied from one person to another but it is a complex network of interactions with other persons and the environment that enables the transfer of knowledge, which includes abstraction. This means that it is essential that the system is learning on its own so that the acquired knowledge is compatible to the already existing experience. Simple implementation of predefined knowledge does not meet the requirements in this respect but would only lead to conflicts.

Conclusion

If computational approaches to mind are to be successful, they must include the ability to recognize the salience, significance and relevance of elements of a perceptual context that are meaningful. Recognition of symbols is insufficient. The successful computational mind must be able to engage in abstraction and meta-abstraction including self-awareness. It must be able to imagine, to volitionally blend abstractions and elements of abstractions in novel ways that allow it to recognize different gestalts in context rather than a series of distinct symbolic elements. And it must be able to discern the validity of these blendings in ways that, in humans, arise from a *sensus communis*.

We conclude that the mind is an emergent phenomenon that is grounded in the brain and influenced by its functions. Abstraction is an emergent capability of the brain, so that it cannot be reduced to physical functions. The emergent qualities of mind include the qualities of consciousness, as well as the capacities for feeling, imagination and volition as which they become present to the mind as part of the meta-abstraction of self-awareness. We find abstraction and logic as prominent features of the mind that must be considered in order to move towards a more viable computational comprehension of the human mind.

References

Agre, P. E. (1988). *The Dynamic Structure of Everyday Life* (MIT AI Technical Report, October 1988, No. 1085). Cambridge, MA: MIT.

Anderson, R. C., Spiro, R. J., & Montague, W. E. (Eds.). (1984). *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum.

Barbiero, D. (2004). Tacit knowledge [Electronic Version]. *Dictionary of Philosophy of Mind*. Retrieved 2009.09.10 from http://philosophy.uwaterloo.ca/MindDict.

Brooks, R. A. (1988). Intelligence without Representation. In J. Haugeland (Ed.), *Mind Design*. Cambridge, MA: The MIT Press.
D'Andrade, R. (1995). *The Development of Cognitive Anthropology*. Cambridge: Cambridge University Press.
Davis, P. M. (1991). *Cognition and learning: A review of the literature with reference to ethnolinguistic minorities*. Dallas, TX: Summer Institute of Linguistics.
DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology, 23*, 263-288.
Dreyfus, H. L. (1972). *What computers can't do: a critique of artificial reason* (1st ed.). New York: Harper & Row.
Dreyfus, H. L. (2007). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. *Artificial Intelligence, 171*(2007), 1137-1160.
Fisher, G. H. (1967). Measuring Ambiguity. *American Journal of Psychology, 80*, 541-547.
Freeman, W. J. (1991). The physiology of perception. *Scientific American, 242*.
Freeman, W. J. (1995). *Societies of Brains: A study in the neuroscience of love and hate, The Spinoza Lectures, Amsterdam, The Netherlands* (Vol. 59). Hillsdale, NJ: Lawrence Erlbaum Associates.
Gadamer, H.-G. (1975 revised trans. by J. Weinsheimer and D.G. Marshall). *Truth and Method* (Second, Revised ed.). New York: Continuum.
Hagengruber, R., & Riss, U. V. (2007). Knowledge in Action. In G. Dodig-Crnkovic (Ed.), *Computation, Information, Cognition - The Nexus and The Liminal* (pp. 134-146). Cambridge, UK: Cambridge Scholars Publishing.
Haugeland, J. (1998). *Having Thought: Essays in the Metaphysics of Mind*. Cambridge, MA: Harvard University Press.
Heidegger, M. (1927). *Being and Time* (J. Macquarrie, E. Robinson, 1962 trans. ed.). New York: Harper and Row.
Honderich, T. (1995). *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.
Keefe, R. (2000). *Theories of Vagueness*. Cambridge, UK: Cambridge University Press.
Kern, A. (2006). *Quellen des Wissens*. Frankfurt: Suhrkamp.
Mai, H. (2009). *Michael Polanyis Fundamentalphilosophie*. Freiburg: Karl Alber.
Merleau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge.
Mingers, J. (2001). Embodying information systems: the contribution of phenomenology. *Information and Organization, 11*(2), 103-128.
Newell, A., & Simon, H. A. (1988). Computer Science as Empirical Inquiry: Symbols and Search. In J. Haugeland (Ed.), *Mind Design*. Cambridge, MA: MIT Press.
Polanyi, M. (1962). *Personal Knowledge*. Chicago: The University of Chicago Press.
Polanyi, M. (1969). Knowing and Being. In M. Grene (Ed.), *Knowing and Being: Essays by M. Polanyi*. Chicago: The University of Chicago Press.
Ruben, P. (1978). *Dialektik und Arbeit der Philosophie*. Köln: Pahl-Rugenstein.
Rumelhart, D. E., & McClelland, J. L. (1986). Parallel Distributed Processing: Exploration in the microstructure of cognition, Vols. 1 & 2. In

Psychological and Biological Models (Vol. 1 & 2). Cambridge: The MIT Press.
Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*, 379-423 and 623-656.
Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
Shapiro, S. (2000). Classical Logic [Electronic Version]. *Stanford Encyclopedia of Philosophy*. Retrieved 2009.09.10 from http://plato.stanford.edu/entries/logic-classical/.
Strauss, C., & Quinn, N. (1997). *A cognitive theory of cultural meaning*. Cambridge: Cambridge University Press.
van Gelder, T. (1997). Dynamics and cognition. In J. Haugeland (Ed.), *Mind Design II* (A Bradford Book ed.). Cambridge, MA: The MIT Press.
Walsham, G. (2005). Knowledge Management Systems: Representation and Communication in Context. *Systems, Signs & Action, 1*(1), 6-18.
Winograd, T., & Flores, F. (1987). *Understanding Computers and Cognition: A New Foundation for Design*. Boston: Addison-Wesley.
Wittgenstein, L. (1953). *Philosophical Investigations* (G.E.M. Anscombe 1962 trans. ed. ed.). Oxford: Blackwell.

Key Terms and Their Definitions

Abstraction

The transition from the practical or theoretical treatment of different but similar (equivalent) concrete objects (or ensembles of objects) of a given domain to the practical to theoretical treatments of these objects (or ensembles of objects) as representatives of that quality with respect to which the equality has been asserted. Consequently abstraction can be seen as the transition from the consideration of equivalent concrete objects to the consideration of the class to which these objects belong regarding their equivalence (adapted from Ruben, 1978).

Logic

Logic, or better formal logic, is the theory of sound reasoning, governed by well-defined rules. It is expressed in mathematical or algorithmic systems, which derive from the sequential application of the principle rules to symbolic expressions, forming a formal language. These deductive systems capture, codify, or record inferences that are correct within the given formal language. One of the fundamental principles of traditional formal logic is the Law of Bivalence, i.e., that a meaningful proposition formulated in this formal language is either true or false. The limitations of such formal systems appear for example in Gödel's theorem (adapted from Honderich, 1995; Shapiro, 2000).

Capacity

Capacity means an ability or power of a thing or person. It can be innate or acquired and describes a causally effective feature of an object.

Examples for capacities are the property of wood to burn or the property of a person to be able to drive a car. To say that an object possesses a capacity does not include that this is true under arbitrary circumstances such as wet wood which might not burn. (adapted from Honderich, 1995).

Embodied

Embodied refers to the integrated nature of cognition with our physical body. It is a recognition of the inseparable nature of 'mind' from 'body' and 'mind' as consisting of more than rational and logical capacities, including emotional, motivational, and experiential capacities.

Embedded

Embedded refers to the relationship between embodied human experience and the world. It is the recognition that immersion in the world is an inescapable fact of human existence, and that the world in which we are embedded consists of not only the physical world but also the cultural and contextualized understanding that we create for it.

Schemas

Schemas are patterns of strongly connected elements of cognition that activate based on salient elements of a particular context and serve as auto-completion processors, allowing us to perceive a gestalt. As strongly connected networks within cognition, they have a bias in activation through repeated exposure to the same or similar stimulus, but they are not rigid and inflexible. They are adaptable, sometimes resulting in the strengthening of existing schemas, sometimes in their weakening in the face of new experience (adapted from D'Andrade, 1995; Strauss & Quinn, 1997).

Tacit Knowledge

Tacit knowledge characterizes a person's capacity to act, to abstract, to make judgments, and so forth without explicit reflection on principles or rules. The person's action is not based on a theory of his or her doing; he or she just performs skillfully without deliberation (adapted from Barbiero, 2004).

Subsidiary awareness

Subsidiary awareness describes that that an object is recognized as part of a gestalt. This means that it is not in the center of the person's attention but on inquiry the respective person is able to identify the particular object as part of the gestalt (adapted from Mai, 2009; Polanyi, 1962).

Focal awareness

Focal awareness describes that an object attracts the attention of a person in contrast to subsidiary awareness. Focal awareness is directed towards the objects of a persons' current interest or activity (adapted from Mai, 2009; Polanyi, 1962).

Ready-to-hand

Ready-to-hand refers to the holistic affordability for action that surrounds an object rather than discrete characteristics or qualities of an object (adapted from Heidegger, 1927).

Present-at-hand

Present-at-hand refers to the salient element, feature or phenomenon that holds the focus of our attention and, because of the temporal nature of our being-in-the-world, is continuously shifting from one thing to another and constrained by *ready-to-hand* possibilities (adapted from Heidegger, 1927).

Sensus communis

Sensus communis is the shared, cultural understanding we create as an essential part of the sense making in which we engage as part of our experience. We use this term in the Gadamerian sense (Gadamer 1975)—the whole set of unstated assumptions, prejudices, and values that are taken for granted; the non-reflective judgments and values learned but not judged.