Self-organization of complex, intelligent systems: an action ontology for transdisciplinary integration

Francis Heylighen

Center Leo Apostel, Vrije Universiteit Brussel, Krijgskundestraat 33, 1160 Brussel, Belgium

tel +32-2-640 67 37, fheyligh@vub.ac.be

Abstract: this paper reviews the general philosophy underlying the transdisciplinary research in the Evolution, Complexity and Cognition (ECCO) group. The ECCO conceptual framework is based on an ontology of action: the fundamental constituents of reality are seen as actions and the agents that produce them. More complex phenomena are conceived as self-organizing networks of interacting agents that evolve to become increasingly complex, adaptive and intelligent systems. The resulting worldview allows us to address the most fundamental issues of philosophy, including metaphysics, epistemology, ethics, futurology and praxeology. It in particular tackles the recurrent issues surrounding the matter-mind duality, including the origins of purposefulness and of subjective experience, and the relation between first-person and third-person perspectives. It achieves this by extending the intentional stance down to the simplest agents, elementary particles. This action-based view moreover supports a variety of practical applications, including the design of self-organizing technological systems, of systems that mobilize people to work in a motivated and coordinated manner, and of systems that support the collaborative development and dissemination of knowledge networks. The appendix of the paper, which is structured as a glossary, systematically defines and surveys the fundamental concepts of the ECCO framework.

Keywords: evolution, complexity, cognition, self-organization, transdisciplinarity, action ontology, systems theory, cybernetics, agents, matter-mind dualism, worldview.

Author biography: Francis Heylighen is a research professor affiliated with the interdisciplinary Center Leo Apostel at the Vrije Universiteit Brussel, and the present director of Evolution, Complexity and Cognition research group there. He researches the self-organization and evolution of complex, social and cognitive systems, which he approaches from a cybernetic perspective. Heylighen has authored over 100 scientific publications in a variety of disciplines, including a monograph and four edited books. Since 1990 he is an editor of the Principia Cybernetica Project, an international organization devoted to the computer-supported collaborative development of an interdisciplinary knowledge network.

Statement: the fully completed Special Issue Author Submission Checklist accompanies the manuscript; the final submission has not been previously published and is not under consideration for publication elsewhere in either print or electronic form.

Introduction

This paper sketches the general philosophy underlying the research in the *Evolution, Complexity and Cognition* (ECCO) group (ecco.vub.ac.be) that I direct. This research group grew out of the international *Principia Cybernetica Project* (PCP) and the *Center Leo Apostel* (CLEA) at the *Vrije Universiteit Brussel*—with which ECCO remains affiliated. From PCP [Heylighen, Joslyn and Turchin, 1991; Heylighen, 2000; pcp.vub.ac.be], ECCO inherited most of its evolutionary-cybernetic philosophy [Turchin, 1977], and its methodology based on the collaborative development of a knowledge network via the web. From CLEA, ECCO adopted the general aim of building a coherent worldview as a remedy against the on-going fragmentation in science and society [Aerts et al., 2002]. Like both CLEA and PCP, ECCO aims at *transdisciplinary integration*, i.e. at the development of a unified conceptual framework that can be applied to problems in all the scientific and cultural disciplines, from the natural sciences via engineering and the social sciences to the humanities.

As our name implies, we find the foundations for this framework at the point where the three approaches of *complexity*, *evolution*, and *cognition* meet. The present paper mostly discusses my own view of how to formulate these foundations, albeit influenced by the ideas of many other ECCO members, including John Stewart [2000, 2008], Carlos Gershenson [2007], Marko Rodriguez [e.g. Rodriguez & Watkins, 2009], Iavor Kostov, Clement Vidal [2008], Mixel Kiemen [2006], Jan Bernheim [1999], Nagarjuna G. [Kharatmal & Nagarjuna, 2008], and Bertin Martens [2004]. While these colleagues are unlikely to agree with everything I write down here, I hope they will broadly agree with the general approach I sketch.

The emerging science of complex systems extends the tradition of *general systems theory* [von Bertalanffy, 1973; Boulding, 1956], which sought to unify science by uncovering the principles common to the holistic organization of all systems, from molecules and cells to minds and societies. However, the classical systems approach had two major shortcomings: the systems it studied were considered as (1) well-defined, static structures, (2) which are objectively given. These assumptions simply do not work for complex adaptive systems, such as societies, minds, or markets [Holland, 1996; Axelrod & Cohen, 1999]. In these systems, structures tend to be fuzzy, variable and to an important degree subjective [Gershenson & Heylighen, 2004; Heylighen, Cilliers & Gershenson, 2007]: different observers will typically distinguish or emphasize different components, boundaries or relationships.

To really understand systems, you need to know how they come into being and gradually develop some form of structure. This brings us to the second strand of our conceptual framework: *self-organization and evolution*. Self-organization is the spontaneous process through which systems emerge and evolve, becoming ever more complex, more adaptive, and more synergetic [Heylighen, 2002]. We see self-organization as the mutual adaptation and co-evolution of the system's initially autonomous components, the *agents*. Agents can be molecules, cells, organisms or organizations. Through their interactions, agents develop a network of increasingly synergetic relationships that coordinate their activities [Heylighen, 2008]. Through continuing evolution based on variation and selection internal to the system, this system becomes ever more complex, more adaptive, and more synergetic. Evolution in the traditional, Darwinian sense is then merely the adaptation of the system as a whole to its encompassing environment, driven by external, or "natural", selection. This holistic view of self-organization and evolution as two aspects of the same process of spontaneous adaptation [Heylighen, 1999a, 2007b] allows us to overcome the pitfalls of genetic or biological reductionism that are often associated with Darwinian approaches.

This coordination between the agents pools their resources, material as well as informational, so that the group as a whole can act more effectively and intelligently than each agent individually. This is the origin of *collective intelligence* [Heylighen, 1999b]. Moreover, the system further increases its knowledge and intelligence by interacting with its environment, as it learns from these interactions so as to become ever more effective in anticipating phenomena and choosing appropriate actions [Heylighen, 2007c]. Thus, the network of relationships between the agents starts to behave like a neural network, capable of increasingly sophisticated cognitive processes [Heylighen, Heath & Van Overwalle, 2004]. Note, though, that neural networks, and more generally self-organizing systems, are intrinsically *distributed* in their organization: it is in general difficult to distinguish separate components or subsystems performing separate functions; the components cooperate as a whole.

The second shortcoming of classical systems theory, its assumption of objectivity, is overcome by noting that knowledge cannot be developed through passive observation of what “objectively” exists, but only through active construction combining a variety of subjective experiences. This leads us into the domain of cognitive science [Thagard, 1996], which until recently was also stifled by a too reductionist and static perspective. The newer approaches, however, emphasize the constant evolution and self-organization of knowledge, and the on-going interactions between subject and environment [Varela, Thompson & Rosch, 1992; Heylighen, 2007d]. This helps us to understand the intrinsic limitations, subjectivity and context-dependence of models [Gershenson & Heylighen, 2004], while still providing us with heuristics to improve our knowledge—however subjective or limited.

The integration of these three perspectives—cognition, complex systems, and self-organizing evolution—points us to a wholly new philosophy of nature, mind and society. It sees the essential building blocks of the universe as actions and interactions, rather than as pieces of matter or energy. Their most important product is intelligent organization, which can be found at all levels, from molecules to global society. For us, this deep metaphysical perspective is at the same time a starting point for concrete, scientific research with plenty of practical applications: we develop and test our fundamental theories by applying them directly to concrete problems.

The problems that presently confront individuals, organizations and society at large all concern complex, evolving systems, such as the global ecosystem, society, the economy, and our own internal system of thoughts and emotions. Thanks to the successes of classical, reductionist science, most of the simple problems have already been solved. The issues that remain are typically ill defined, open-ended, constantly changing, and with ramifications extending into an unlimited number of other domains. Coping with these problems requires a set of new methods that take complexity and change as their starting points [Battram, 1996; Axelrod & Cohen, 1999].

The advantage of the ECCO approach, with its high level of generality, is that the concepts it produces are applicable to any system from any domain—whether biological, technological, mental or social. These concepts are applicable in particular to hybrid or mixed systems, such as the World-Wide Web with its technological, social, economical and psychological aspects [Heylighen, 2007a,b,c,], or such as social software used to support innovation in organizations [Kiemen et al., 2009].

Unlike other high-level, abstract approaches, our concepts directly address real-world problems and the tools to tackle them. Indeed, we see *evolution* as a giant problem-solving process in which systems are constantly trying to adapt to new circumstances, or improve their handling of existing situations [Heylighen, 2007c]. *Cognition* is merely an interiorization of this on-going process of trial-and-error and a registration of shortcuts that have proven to be useful for re-application later [Campbell, 1974]. *Complexity* is both a feature of the problems that need to be solved, and of the solutions that are most robust in handling multifarious and ever-changing demands [Axelrod & Cohen, 1999].

Therefore, the ECCO perspective encompasses both the most abstract realms of ontology, epistemology and metaphysics, and the most concrete methods to solve problems in everyday life, organizations, technology, and society. These two aspects constantly interact and feed back into each other: practical experience in tackling problems suggests new concepts and principles for understanding complexity in science and philosophy. Clarifications and integrations in our theoretical framework, on the other hand, immediately suggest new ways to tackle concrete problems. For example, the concept of *stigmergy* [Heylighen, 2007a] proposes both a foundation for the theory of self-organization, and a practical method for creating effective web collaboration systems, such as Wikipedia.

Thus, our most important results to date are twofold:

1) on the theoretical level: a coherent and comprehensive philosophy or worldview, including ontology, metaphysics, epistemology, futurology, axiology/ethics, and praxeology, based on these ideas;
2) on the practical level: a collection of methods and technologies to support self-organization and collective intelligence, and in particular the self-organizing interaction between people who together develop complex knowledge systems.

I will now present these results in more detail, starting with a summary of the worldview. I will then go into more detail about the ontological and epistemological foundations of our worldview, by examining the implications of using *action* as the primitive of our conceptual framework. I will in particular show how an ontology based on action avoids the problems arising from the opposition that is traditionally conceived between mind and matter (dualism), and between subject and object (first-person vs. third-person perspectives). I will then review our "praxeology", i.e. a number of practical methods and applications entailed by our action-based framework. I will in particular focus on how these practical methods help us to further develop our own theory, as well as knowledge systems in general. The appendix to the paper will review the whole conceptual framework in a more systematic way, using a glossary format in order to define increasingly complex concepts starting from the single primitive of "action".

The ECCO worldview

For centuries, people have been wondering about their existence and place in the universe. The fundamental questions they have been asking can be classified in the following categories, each defining a fundamental philosophical domain:

1. *What exists? What is the ultimate reality?*

    Ontology: defining the constituents of reality

2. *Why is the world the way it is? Where do we come from? Why is there something rather than nothing?*

    Metaphysics: determining the origins or ultimate causes

3. *Where are we going? What is the ultimate fate of the universe?*

    Futurology: forecasting the future

4. *What is good and what is evil? What should we strive for? What is the meaning of life?*

    Axiology: a system of goals, values, and ethics

5. *How should we act? How can we tackle our problems?*

    Praxeology: guidelines for practical action

6. W*hat is true and what is false? How can we acquire reliable knowledge?*

    Epistemology: a theory of knowledge

The answers to all these questions together determine an integrative *worldview*, i.e. a comprehensive philosophical system, a coherent vision of the whole [Vidal, 2009; Heylighen, 2000; Aerts et al., 2002]. A worldview gives meaning to our life, and helps us to understand the world around us. A coherent worldview is particularly important in the current era of accelerating scientific, cultural and social developments, in which all the old certainties are put into question. The confusion and fragmentation associated with this contribute to alienation, pessimism and uncertainty [Geyer, 1994], and the need for psychological guidance in the form of a clear and reliable system of thought.

Unfortunately such a framework is all too often found in fundamentalist ideologies, or in irrational beliefs and superstitions. Science should be our weapon in the fight against irrationality and fundamentalism. Regrettably, contemporary science seems to contribute to the confusion by the avalanche of often-contradictory observations and theories that it overloads us with. That is why we need to develop a coherent, new worldview that is solidly rooted in the most advanced scientific concepts and observations, but that goes beyond the simple-minded determinism and reductionism of Newtonian science in order to bring us back a sense of meaning, of purpose, of being part of a larger whole.

Our ECCO philosophy tries to show how the different scientific and philosophical insights can be integrated in a coherent framework. This framework is based on the process of evolution as a spontaneous force or drive for the self-organization of increasingly complex and intelligent systems. This process can be analyzed into elementary transitions or changes that we call "actions". This evolutionary process leads from particles to atoms, molecules, cells, organisms, humans, and societies to the emerging "global brain" [Heylighen, 2007bc]. Let us summarize this philosophy by the way its answers the fundamental questions:

Ontology

> the most fundamental components of reality are actions and agents. These are elementary processes or transitions between states, not independent, timeless pieces of matter. Therefore, our ontology is fundamentally holistic and dynamic, not static and reductionistic. Out of their interactions, organization emerges. As these organizations or systems become more complex and adaptive [Heylighen, 1999b], they start to exhibit increasingly sophisticated forms of cognition or intelligence, i.e. the ability to make informed choices between actions. This ontology rejects the traditional assumption of *dualism*, which sees matter and mind as the two fundamental but independent constituents of reality: both matter and mind are merely aspects of the same underlying network of actions.

Metaphysics:

> if we go could go back in time, towards the origin of the universe, we would see how agents and systems become ever simpler, until they lose any form of complexity or organization. The organization we see around us now can be explained by the processes of blind variation, that have been producing random combinations of agents and actions, and natural selection, that have retained only those combinations that are "fit", i.e. adapted internally to each other, and externally to their encompassing environment. Since natural selection or self-organization is a spontaneous, automatic process, there is no need to postulate external or supernatural causes, such as God or the Laws of Nature, to explain the origin of the phenomena we see around us [Heylighen, 2000, 2010].

Futurology:

> this process of on-going complexification and adaptation can be extrapolated towards the future. This allows us to predict that in the medium term conflict and friction within human society will diminish, cooperation will expand to the planetary level, well-being will increase, individuals will become ever more integrated with the socio-technological systems that surround them, while individual and collective intelligence will spectacularly augment [Heylighen, 2007b,c; Stewart, 2000]. In the long term, this increase in cooperation and evolvability is likely to expand beyond the planet into the universe [Stewart, 2010]. However, since evolution is an often chaotic process of trial-and-error that is not accurately predictable, we should be ready for various unexpected problems and setbacks along the road.

Axiology:

> the inner drive or implicit value governing all life is *fitness*, i.e. survival, growth and development. In the present human situation, this fundamental value can be translated as a universal and sustainable *quality-of-life*, *well-being* or *happiness* [Heylighen & Bernheim, 2000]. Evolutionary, psychological, and cybernetic theories allow us to derive a number of more concrete values from this overarching value, i.e. properties that are necessary for long-term well-being. These include openness, diversity, intelligence, knowledge, cooperation, freedom, personal control, health, and a coherent worldview. In the longer term, fitness implies increasing adaptivity and evolvability beyond human society as we know it. Actions that promote these values are intrinsically good; actions that suppress them are bad.

Praxeology:

> to maximally achieve these values in real life, we will need to overcome a variety of problems and obstacles. Cognitive science, cybernetics, and complex systems science suggest various tools and strategies to tackle complex problems [e.g. Heylighen & Vidal, 2008], and to stimulate and steer self-organization so as to be as efficient as possible. These methods include feedback, anticipation, hierarchical decomposition, heuristic search, stigmergic coordination, and memetic engineering. At the level of society, these methods define a strategy for effective governance, for the maximization of collective intelligence, and the minimization of friction and conflicts.

Epistemology:

> in order to solve problems, we need adequate knowledge. Knowledge is not an objective reflection of reality, though, but a simple model that makes useful predictions. Different problems may require different models of the same reality, without any one being the "true" representation. However, models that make more wide-ranging, accurate and reliable predictions are intrinsically better. Cognitive science, cybernetics, and neuroscience help us to understand how the brain learns from experience and makes predictions via the self-organization of neural patterns, and the feedback between perception and conception, observation and theory [Heylighen, 2007d; Hawkins & Blakeslee, 2005]. Similar mechanisms may be implemented as computer algorithms to extract new knowledge from unstructured data, and thus discover better concepts and theories.

A non-dualist ontology of action

After this bird's eye view of the ECCO philosophy, let me try to explain the fundamental ideas behind this perspective in more depth. These concepts start from the notion of *action* as the most primitive element or unit of reality. An action is an elementary process or, more simply, a *change*: a transition from an initial state of affairs (cause) to a subsequent one (effect). However, our conceptual framework avoids reducing such a dynamic "action" to the static concept of "state" by defining a *state* itself as the collection of actions that are possible in that state [cf. Turchin, 1993]. Thus, states are defined by actions, while actions are defined in terms of states. This is a recursive or bootstrapping definition that allows us to avoid postulating absolutist foundations [Heylighen, 1990, 2001]. An *agent* can then be defined as an aspect or part of a state that is necessary for the action to occur, but that persists during the subsequent change. Thus, an agent can be seen as a cause or producer of actions that does not vanish after the action. Particles, objects, or pieces of matter—the static constituents that form the basis of the more traditional scientific worldview—are then merely simple types of agents. For a more precise definition of "action", "state", "agent" and all the more advanced concepts derived from these, I refer to the appendix of this paper.

This action philosophy may be situated within the broad tradition of *process metaphysics*, which considers change ("becoming") as more fundamental than static existence ("being") [Heylighen, 1990b; Prigogine, 1980]. This perspective goes back to the pre-Socratic philosopher Heraclitus and to the Chinese philosophy of the Tao. However, it has largely disappeared from the Western worldview after Parmenides, Plato, and Newton founded their ontologies on eternal, unchanging constituents: abstract ideas, objects, matter, space, and the laws of nature. In contrast to earlier process philosophies [e.g. Whitehead, 1978; Rescher, 1996], which tend to be vague, difficult to grasp and somewhat mystical, the action approach is concrete and practical, as it is concerned with the actions that we as subjects perform in the real world.

Since an action normally is followed by—or elicits—subsequent actions, actions are intrinsically connected, forming a complex network of causes and effects. Within this network, higher-order structures can be distinguished, including systems of interacting agents. Actions are assumed to be intrinsically *directed* or oriented, from cause to effect, or from past to future. As such they are in general irreversible. This directionality is inherited by the producers of action: agents and systems of agents are assumed to have implicit *preferences* that give a direction to their further evolution [Heylighen, 1999a; Stewart, 2000]. At the level of individual agents, this preference translates into the intuitively "mental" notions of *goals* or *values*.

Moreover, an action is normally always a response or "re-action" to the particular state of affairs or condition that functions as the action's cause. This implies that agents are intrinsically *sensitive* to outside conditions: they sense or "perceive" specific conditions in their environment, acting differently under different conditions. Since, in accordance with the uncertainty principle of quantum mechanics, we do not assume determinism, agents have in general more than one way to react to a certain condition. This brings us to a rudimentary notion of *freedom* or *choice* [Turchin, 1993]. If the agent systematically tends to make "good" choices from among the possible actions (in the sense of actions that bring the agent closer to its goals), we will be tempted to attribute a rudimentary form of *knowledge* or *intelligence* to that agent. The "freedom" of a particle or molecule is of course very limited, and its "decision-making" function appears either pre-determined or random. On the other hand, more complex agents, such as bacteria, organisms and people, have evolved increasingly more sophisticated decision-making mechanisms, and this is, of course, the basis of cognition.

Thus, it becomes clear how the notions of action and agent, which are defined in such a way that they encompass the simplest possible particles, very quickly lead us to consider "mental" attributes, such as goals, sensations, choices, knowledge and intelligence. This allows us to immediately transcend the mind-matter duality. Our action ontology sees matter and mind as merely complementary aspects of the behavior of agents, respectively the *causal* (past-determined) and the *intentional* (future-directed) aspects. Practically, this means that we can describe particles, objects, cells, animals, humans and societies on the basis of the same conceptual framework.

This is particularly important for transdisciplinary integration, which requires bridging the gap between the sciences of matter (physics, chemistry, physiology, etc.) and the sciences of mind (psychology, sociology, philosophy, history, literature, etc.). The physical sciences typically start from a mechanistic, materialistic ontology, in which all phenomena are reduced to combinations of material particles that obey the deterministic, absolute laws of causality. In such a Newtonian worldview, if you know the initial state of a particle or system and the forces that impinge upon it, then you can perfectly predict the further trajectory of the system [Heylighen, 1990b, 1989]. This perspective may be called the *causal stance*. From this perspective, the cause (initial state) is sufficient to determine the effect (all future states).

The sciences of mind, on the other hand, typically start from the notion of an intelligent agent who has *beliefs* about the state in which s/he lives, *desires* about what that state ideally ought to be, and *intentions*, in the sense of actions that the agent intends to execute in order to bring the present state closer to the desired state. This perspective may be called the *intentional stance* [Dennett, 1989]. It assumes that the behavior of an agent can best be predicted by starting from the agent's preference or goal, and the agent's belief about its present conditions. Preferences and beliefs together allow you to infer the agent's intentions, and therefore its actions.

As Dennett [1989] has argued, intentional and causal (or what he calls "physical") stances are not contradictory theories of the world: they are merely different perspectives or approaches to analyze, model and predict phenomena. The causal stance typically works better for simple, deterministic systems where the observer knows everything about the things that will affect the system, so that predictions are reliable. The intentional stance works better for complex systems with a lot of uncertainty about further events. In the latter case, the precise sequence of actions (i.e. the trajectory that the agent follows through its space of states) is in practice impossible to predict. However, the general outcome may well be predictable, in the sense that the agent eventually will move closer to its goal, by counteracting any unexpected disturbances that make it deviate from its preference. This phenomenon of convergence towards a "preferred" state has been called *equifinalit*y [von Bertalanffy, 1973]: different initial states tend to lead to the same final state(s) (in the terminology of complex systems, this set of final states is called the "attractor" of the dynamics [Heylighen, 2002]). At a deeper, formal level, causal and intentional stances are equivalent, since they both can be expressed in a mathematical framework where a system follows a trajectory through its state space that maximizes some "preference" criterion [Heylighen, 1990b].

Our action ontology starts from the intentional stance, because it is intrinsically more flexible than the causal stance, and can be applied in more complex circumstances. Therefore, we consider all systems—whether "physical" or "mental"—as intentional, sensing agents. Some philosophers may interpret this as a form of *panpsychism* [Nagel, 1979], i.e. the attribution of mental properties to all phenomena, whether they are particles, cells, or people.

But our intent is not to reduce everything to the elusive notion of mind, but to transcend the mind-matter duality by starting from concepts that are as simple and operational as possible. For example, the "intention" of a molecule is nothing more than its tendency to move along a gradient towards a state of minimal potential energy, and its "sensation" is nothing more than its sensitivity to changes in initial conditions. From the point of view of the ontology of action, material particles and human minds alike can be conceptualized in the same basic terms, without any loss of predictive power.

Another possible interpretation of our philosophy is *animism*, i.e. the belief—typical of "primitive" cultures of hunter-gatherers—that all phenomena, such as trees, animals, or mountains, are sentient beings. The advantage of an animist worldview is that it a priori avoids *alienation* [Charlton, 2002, 2007], i.e. the feeling that we do not really belong to the environment that surrounds us. For the alienated person of the industrial age, the natural environment consists of impersonal, foreign phenomena, to be either feared and suppressed, or manipulated and exploited. For an animist, on the other hand, these phenomena are beings to interact with on an equal footing—as potential allies, rivals or enemies, but never as cold, impassive "things".

Animism has been nearly universally rejected as naïve, because it anthropomorphizes simple phenomena into human-like intelligences. But the intentional stance or agent ontology does not presuppose any near-human level of intelligence: it merely attributes to all agents in-built preferences, the ability to sense certain conditions, and the tendency to react to these conditions by appropriate actions. These minimal assumptions apply equally well to elementary particles and to intelligent human beings. As such, they restore a continuity and interactivity to the world that prevent us from feeling alienated from nature.

Of course, these different agents differ radically in their level of complexity or organization. As agents become more complex and intelligent, they start to exhibit more advanced "mental" qualities, such as memory, feeling, emotion, rationality or consciousness. But our underlying philosophy sees this evolution as continuous. It does not presuppose any strict boundaries between systems that exhibit these qualities (e.g. humans and higher animals) and systems that do not (e.g. insects, plants or rocks). That makes it much easier to understand the origin of complex and mysterious phenomena, such as consciousness or intelligence, by retracing their origin in much simpler phenomena.

To explain this growth of complexity [Heylighen, 1999a], we need to study the collective level that appears when several agents interact. The intrinsic directionality of action will here lead to self-organization, a spontaneous evolution towards increasingly effective coordination and cooperation between individual actions. This process constantly generates more complex and intelligent behavior. Since I have explained the mechanisms and implications of self-organization in detail in earlier papers [e.g. Heylighen, 2002, 2008], and since its basic concepts are recapitulated in the appendix, I will here not go more deeply into this issue. Instead, I will further investigate the "mental" aspects of our action philosophy, before reviewing its practical applications.

Subjective experience and the first-person perspective

One of the most fundamental problems of the Newtonian, materialist worldview appears to be its inability to explain consciousness, or what may be termed more accurately *subjective experience*. This is what a human subject, such as you or I, feels "inside our mind". While this feeling usually, but not always, seems to be triggered by some outside phenomenon, it is intrinsically personal, subjective, emotional, and impossible to accurately convey in words or other symbols. Philosophers sometimes use the term "qualia" to refer to these internally experienced phenomena [Tye, 1986; Solms & Turnbull, 2002].

The fundamental issue is often formulated like what the philosopher Chalmers [1995] has called "the hard problem of consciousness": while materialist scientific theories may be able to explain how a particular outside phenomenon (e.g. a ray of sunlight impinging on the retina in your eye) triggers a process in your brain (e.g. a train of electrical activity traveling from neuron to neuron) which itself triggers a physical reaction (e.g. moving your hand up to protect your eye), these theories do not explain why this chain of cause-and-effect transitions is accompanied by a subjective experience (e.g. the feeling of being blinded). Chalmers imagines a robot-like creature, which he calls a "zombie", that performs all the right actions in reaction to the right stimuli, following a purely causal, mechanical logic, but that does not "feel" anything while acting like this. He concludes that since traditional science cannot distinguish between the behavior of a zombie and the one of a conscious, feeling human being, science is intrinsically unable to solve the problem of consciousness. The only solution he can imagine is to replace materialism by some form of panpsychism, where even elementary particles can "feel" phenomena to some degree, and human beings are merely extremely complex and effective "accumulators" of such elementary feelings.

Other philosophers consider that the problem of subjective experience cannot be tackled by traditional science because science by definition considers phenomena from a third-person perspective, as things "out there" about which you think and reason as if they exist wholly independently of yourself. Instead, they propose that you should examine these phenomena from a *first-person perspective*, as things that you directly, personally experience. This is the approach of *phenomenology*, which tries to analyze phenomena as they are experienced by the subject, by the "I" [Thompson, 2007]. From the first-person perspective, subjective experiences are the only things that count; the "third-person" world of objects and objective relationships between objects, as studied by science, is merely an abstract construction, which is meaningless unless it can be grounded in first-person experiences. Meaning, value, feeling, experience only arise when a phenomenon interacts with the "I", with the first person each of us is. Third-person approaches can only speak about cold, "objective", rational properties, about abstract symbols obeying formal rules, not about purpose and meaning since these emerge only in the experience of the first-person, the subject [Gendlin, 1962; Brier, 2008].

The ontology of action has no difficulty with subjective experience, and therefore it denies that there is an intrinsically "hard" problem of consciousness. First, it is not founded on the existence of independent, material objects obeying objective laws. Therefore, it has no need to reduce notions like purpose, meaning or experience to arrangements of such mechanical entities. Instead, it takes actions as it point of departure. An action, as we defined it, immediately entails the notions of *awareness* or *sensation* (since the agent producing the action needs to sense the situation to which it reacts), of *meaning* (because this sensation has a significance for the agent, namely as the condition that incites a specific action), and of *purpose* (because the action is implicitly directed towards a "goal", which is the attractor of the action dynamics). However, it does this without obscure presuppositions about intangible, supra-physical forces steering the action from the right condition to the right effect: the action is defined simply by the change it produces. Since this change is the most simple, primitive element in our ontology, no further explanation of its features is required [Heylighen, 1990].

Second, if we look more closely, the subjective experiences that form the basis of the phenomenological, first-person perspective are really actions. As the theory of enactive cognition [Varela, Thompson & Rosch, 1992; Thompson, 2005] has argued, all perception is a form of action. We can only see things because our eyes continuously scan the surroundings via saccading movements, so that the image projected on the retina continuously varies. As experiments have shown, if the image on the retina is immobilized, we simply become blind to it. The neurons in our brain stop responding to unchanging stimuli because of the mechanism of "neuronal fatigue": a neuron that is continuously stimulated after a short while loses its ability to transmit activation to other neurons [Menghini et al., 2007]. In other words, for us to become aware of a phenomenon, either the phenomenon needs to change in some way, entailing an action on the part of the phenomenon, or we need to change something in the way we observe the phenomenon (e.g. by turning our head towards it, or scanning it with our eyes), entailing an action on the part of the subject. We may conclude that the most fundamental requirement for something to become an experience, i.e. to reach the awareness of a subject that is sufficiently evolved to have a nervous system, is that it is includes an action.

This requirement can even be generalized to a much simpler agent, such as an elementary particle. Indeed, the "sensation" of such a simple agent is merely the cause to which it responds by producing an effect. Causality is in essence *co-variation*: a change in the initial condition (cause) is followed by a change in the subsequent condition (effect) [Heylighen, 1989]. Without change, there is no (co)variation, therefore no causation, and no action. This is a generalization of Bateson's [2000] famous definition of information as "a difference [initial change] that makes a difference [subsequent change]". Indeed, causation is in essence information transmission, or what I have called "distinction conservation" [Heylighen, 1989]. Information gets its meaning because of its causal import: it influences some agent to act in a way different from the way that agent would have acted without the information. A phenomenon (like Chalmers' hypothetical zombie) that does not produce any difference in any observable action is fundamentally uninformative or meaningless: it might as well not exist.

But less us go back to the action at the origin of *human* experience. This action may be initiated by the subject, or caused by some part of the environment. The difference is fundamental, because it delineates the border between "I" (internal, first person, subject, self) and "it" (external, third person, object, non-self). However, it is not as fundamental as the fact that *an action has taken place*, because the change characterizing an action is the true origin of the sensation. Newborn babies still need to learn to distinguish between "I" and "not-I" [Rochat, 2001]: they may see a movement or hear a noise, but they do not know whether that sensation was caused by their own activity (which may include moving, perceiving, or even imagining), or by something outside themselves. The concept of "object"—and with it the notion of *objectivity*—only arises after the subject has learned to distinguish between purely internal actions (e.g. moving the head in such a way that the image on the retina changes) and actions that affect the outside world (e.g. manipulating a toy, or the toy moving on its own).

An *object* can be defined as a structure that remains invariant under purely internal (i.e. cognitive) actions [Turchin, 1993; von Foerster, 1976]: while the image on the retina changes depending on the perspective, its abstract features (e.g. shape, texture) are not affected by that shift in perspective. This is what distinguishes a sensation caused by an invariant, external object from a sensation with a purely internal origin (e.g. a hallucination, or a piece of dust in the eye). This is also the origin of the distinction we make between imagination and reality: we tend to consider as "real" those phenomena that are invariant across different ways of perceiving them [Heylighen, 1997; Bonsack, 1977].

In this way, the ontology of action makes it in principle possible to start from the first-person world of subjective experiences (phenomenology) and derive from it the third-person world of objects and relationships between objects (science). While it may be useful to distinguish the two perspectives as a way to categorize and integrate different approaches [Wilber, 1997; Combs & Esbjörn-Hargens, 2006], the present philosophy focuses on the *continuity* between these perspectives, considering them as merely aspects of the same ontology of action that can be derived the one from the other. However, we must not forget that this derivation is by definition incomplete, as the subject can never sense all aspects of the external phenomena. Therefore, as the constructivists have argued, our knowledge always remains a preliminary, subjective construction [von Glasersfeld, 1996; Maturana and Varela, 1987]. It can never become an accurate reflection of an objective reality—although we may strive to make it gradually more comprehensive, reliable and less subjective by taking into account ever more diverse perceptions and points of view. In this way, the ontology of action implies an epistemology that is constructivist (all knowledge is intrinsically a subjective construction), but not relativist (some forms of knowledge are more reliable than others) [Heylighen, 1997].

Applications of the ECCO worldview

We noted that the ECCO approach is just as committed to practical applications as to theoretical investigations. This follows from our action-based ontology: if you believe that actions are the most fundamental components of reality, then you will be inclined to invest a lot of energy in making these actions more effective. By clarifying the nature and organization of action, you moreover get an immediate handle on the problem of effectiveness. For example, our concepts of intelligence, synergy, coordination and organization (see appendix) all denote aspects that characterize more or less effective forms of action. By controlling these aspects (e.g. increasing synergy or supporting coordination), you can make your actions more effective.

To clarify the overall strategy, two things should be noted: 1) our philosophy emphasizes the power of self-organization or evolution, i.e. mechanisms that spontaneously, without outside intervention, tend to increase effectiveness; 2) we see evolution as a process that is in general slow, difficult and sometimes painful, as it relies on trial-and-error. This implies many errors, detours and setbacks. The end result of these processes is a form of organization that cuts short across these detours, so that errors do not need to be repeated. Knowing how such organization came into being helps us to accelerate that process, by making sure that all the components that support self-organization are available.

For example, we have observed that stigmergy [Théraulaz & Bonabeau, 1999; Parunak, 2006; see the appendix for a definition] is a very effective mechanism to achieve coordination. Such stigmergy requires a reliable medium via which the agents can indirectly communicate and where they can accumulate the provisional outcomes of their work. Therefore, we can promote coordination between agents by providing them with an adequate medium. A good example of such a medium is a wiki-style website, which can be edited and added to by a variety of independent individuals. Creating such stigmergic interaction medium [Heylighen, 2007a,b] is an example of what we call *design for self-organization*: conceiving and realizing a structure that does not so much tell the agents what they should do, but that helps them find out for themselves what is the most effective way to act in a coordinated, synergetic manner. In self-organization there is no centralized control, no separable agent that directs the actions of the others. Instead, coordinated activity emerges in a distributed manner, out of the tangle of interactions between largely autonomous agents.

The most obvious application domain of this approach is the design of autonomous technologies, i.e. systems that can tackle more or less complex problems without direction or supervision by a human designer, by relying on some form of self-organization [Gershenson, 2007]. Over the past decades, several examples of such systems have been explored and applied in the domain of computing [Heylighen & Gershenson, 2003], including neural networks, genetic algorithms and multi-agent simulations. These self-organizing software systems typically solve problems that are too complex for traditional, centrally controlled methods—such as finding structure in large, noisy and incoherent data sets, or learning how to perform ill-defined tasks such as handwriting recognition.

More recently, self-organization is being applied to engineering problems [Valckenaers et al., 2006; Elmenreich et al., 2009], where hardware systems need to be set up that can cope with unpredictably varying situations [Gershenson & Heylighen, 2004]. Examples include wireless communication networks, where you can never be sure where or when a particular wireless component will be reachable, flexible-manufacturing systems that constantly need to reconfigure themselves in order to meet varying demands, and self-organizing traffic lights that can adapt efficiently to the unpredictable variations in traffic streams [Cools, Gershenson & d'Hooghe, 2007]. These applications typically make use of the mechanisms that are part of our ontology, such as random variation, stigmergy, mediation and the search for synergetic connections. Investigating such practical applications helps us to understand the fundamental mechanisms of self-organization in a more concrete and explicit manner, reducing the risk of getting lost in theoretical abstractions that have little relation with the real world.

When this philosophy is applied to human agents rather than hardware or software components, the support structure we are aiming for may be called a *mobilization system*. Its aim is to encourage and organize people to work towards a common goal. This is particularly useful in our present society that suffers from complexity, fragmentation and information overload, with the result that people have lost a clear sense of goals and values, and often fail to see the forest for the trees. There are presently so many issues that clamor for our attention that we often do not know anymore what to focus on, but instead jump from the one thing to the next without giving anything the attention it deserves, or simply procrastinate and wait for things to be imposed from the outside [Heath & Anderson, 2010; Heylighen & Vidal, 2008]. A mobilization system would combat this confused and unproductive way of acting by redirecting effort in the most efficient way at the most important issues. This requires the following steps:

- helping people to **reach consensus** about the specific goals that they consider most important. This can be done in part by seeking inspiration about fundamental values in the evolutionary worldview [e.g. Heylighen & Bernheim, 2000], in part by creating effective discussion systems that help a group to come to a well-reasoned consensus. Examples of such systems are being developed on the web [Klein, 2007; Malone & Klein, 2007].
- **motivating** and **stimulating** people to work towards the goals that have thus been agreed upon. Here, a very useful paradigm is the concept of "flow" [Csikszentmihalyi, 1990], which specifies the conditions under which people work in the most focused and motivated manner. These conditions are:
    - **clear goals**: there should be minimal ambiguity about what to do next;
    - i**mmediate feedback**: any action should be followed by an easily interpretable result, so that you either get a confirmation that you are on the right track, or a warning that you need to correct your course;
    - **challenges in balance with skills**: tasks should be neither too difficult nor too easy for the people entrusted to perform them, in order to avoid either stress or boredom.

  Additionally, there exists a wide range of techniques from psychology, behavioral economics and memetics that help us to formulate goals and tasks in a way that is maximally motivating, persuasive and easy to follow [Heath & Heath, 2007; Thaler & Sunstein, 2009; Heylighen, 2009]
- **coordinating** and **aggregating** the individual contributions so as to ensure maximum collective results. This can be built on the mechanisms of stigmergy and self-organization mentioned before [Heylighen, 2007a; Parunak, 2006].

By minimizing uncertainty, confusion, friction and procrastination, work that is mobilized by such a system would not only become much more productive and effective, it would also make the participants more satisfied with what they are doing. This mood-enhancing effect can be achieved by reducing their levels of anxiety and boredom, while increasing the satisfaction they get from an activity that proceeds efficiently towards a goal they consider important. The reality of such an effect is confirmed by a variety of empirical studies [e.g. Csikszentmihalyi, 1990; Diener & Biswas-Diener, 2008; Sheldon & Lyubomirsky, 2006], which show how goal-directed and productive engagement with the world is the best recipe for raising happiness. Moreover, happiness itself tends to raise people's productivity [Lyubomirsky et al., 2005], as happy people feel more confident, focused and creative, thus creating a virtuous cycle that further amplifies the positive effect of the envisaged mobilization systems. This fits in with the ECCO worldview, which considers happiness as a good proxy for fitness [Heylighen & Bernheim, 2000], which we see as the most fundamental value driving forward the process of evolution.

More generally, the concept of mobilization system builds further on our ontology, which not only considers the pragmatics aspects of actions, but the subjective experiences that accompany them. Recently, under the influence of particularly inspiring interfaces, such as those offered by Apple's iPhone or Mac computers, the awareness has been growing that a good technology should not only be effective and simple to use, it should offer a memorable *user experience* [Hassenzahl & Tractinsky, 2006]. That means that the user of the technology would develop an intuitive, affective rapport with the system—almost as if it were an extension of the self. A positive user experience is difficult to define, but easy to recognize once you start using the system. Our philosophy, by focusing on the origins of affect and experience in the context of action, provides a good starting point for designing inherently appealing technologies. For example, such technologies should embody the cybernetic criteria of "flow": clear goals, immediate feedback, and challenges in tune with the user's skills.

Self-application of the transdisciplinary worldview

A third application domain is particularly relevant to our own work: the development of complex knowledge systems. We see knowledge as a network of concepts (sometimes called "categories" or "distinctions") that are connected by "if A, then B" rules (represented as A → B) [Heylighen, 2001, 2007d]. Connections denote logical, taxonomical or causal relationships between the concepts, like in the following examples: banana → fruit, banana → sweet, eating → digesting. Concepts are mental representations of what we have defined as states or conditions. The A → B connections between them represent anticipations of B, given A. Both concepts and connections can be fuzzy and variable. Successful rules, i.e. whose predictions are frequently confirmed, are reinforced and thus become better established, while unsuccessful ones are weakened and eventually eliminated [Heylighen, 2007d]. In this way, the knowledge system undergoes a variation and selection type of evolution, which is self-organizing or distributed in the sense that local changes in concepts or connections propagate to neighboring rules, in a sequence of mutual adjustment rounds which may lead to the emergence of new higher order concepts. This representation of knowledge development can in principle be implemented as a computer system that supports the representation, analysis and creation of complex knowledge networks. A few prototypes of such systems have already been implemented within the ECCO community [e.g. Kharatmal & Nagarjuna, 2009; Rodriguez, 2008].

The existence of such tools and methods for knowledge development now becomes particularly useful for our own further research. The ECCO worldview as I have sketched it is a complex conceptual system that is constantly developing in order to become more coherent, more comprehensive and easier to understand and apply. Therefore, it is itself a complex, cognitive system undergoing self-organization or evolution. That means that its own principles should be applicable to understand and steer its own further development. This is a "bootstrapping" approach, where the content of the theory inspires a method to extend the theory, so that theory and method can drive each other forward.

Within ECCO, this idea has been formulated (most actively by Iavor Kostov) as the *stigmergic university* project. A university is an organization whose purpose is to develop and disseminate knowledge, and this in an as wide array of subject domains as possible. These are the functions known respectively as *research* and *teaching* (including *assessment* of progress). The combination of the two is what distinguishes a university from either a research center or a teaching institute. The reason for combining these two functions is that knowledge changes continuously so that everyone, students as well as teachers and researchers, continuously needs to keep up to date. The best people to convey the latest developments are the ones who have themselves been contributing to these developments. Therefore, the research and teaching functions are best combined. However, in the traditional university organization both teaching and research demand a lot of effort, so that the most active researchers rarely find the time to teach, while the most active teachers cannot afford to do much research. This activity could be organized much more efficiently by using the kind of mobilization systems sketched earlier, with a focus on systems that support knowledge structuring.

The "stigmergic university" system we envisage would contain at least the following components;

- a website that represents knowledge in the form of a **conceptual network**. Each concept would be explained by a document containing a definition, further explanation, examples, and links with connected concepts. This was the original conception behind the Principia Cybernetica Project [Heylighen, Joslyn & Turchin, 1991]. Similar to Wikipedia, users of the systems would be able to edit the text and to create new concept pages. Unlike Wikipedia, concepts would be linked as much as possible using formally labeled connections [Nagarjuna, 2007; Heylighen, 2001], so that the knowledge would have the shape of a semantic and associative network rather than an unstructured encyclopedia.
- a set of programs to **analyze the structure** of this formal network. This module would be able to pinpoint ambiguous, missing, inconsistent, or redundant information, and to suggest ways to improve the network, e.g. by proposing additional links, changing the types of links, or new, higher-order categories that cluster a number of related concepts together [Heylighen, 2001]. This is the basic "research support" module, which helps individual researchers to discover and resolve the remaining gaps in their understanding.

- a **discussion module** that would support reaching consensus about as yet unresolved issues, not only via traditional free-form "forums", but via structured networks of arguments and counterarguments [Klein, 2007]. Here, the systems' analysis of the network would help the users to identify the weakest arguments, thus stimulating their authors to either change their position, or propose better arguments to defend it. This module further supports research, but now at the level of the dialogue between researchers who have different perspectives on the issue.
- an educational subnetwork of "**guided tours**" or paths within the conceptual network, functioning as suggested sequences for learners to study the knowledge that start from easier concepts and build up towards more advanced ones depending on the interests of the learner. This would help a student to quickly advance into a complex knowledge domain without getting confused by the variety and complexity of the knowledge that is available. To make this most effective, the guidance should be adapted individually to the interests and skills of each student. This can be achieved with the following assessment module. Moreover, the system would learn from the way it is being used, so that later students would get recommendations that take into account the experiences of previous students.
- a system that **generates questions** about the knowledge in the network. Most of these questions should be easy to answer by someone who knows the domain, and thus could be used to test the knowledge of the users. This is the module for *teaching and assessment* on the basis of "inverted learning" [Martin, 2009]. The procedure starts by asking the users a question in order to stimulate or *challenge* them to reflect about the domain. It then gives them the right answer on the basis of what is contained in the knowledge base (*feedback*). Depending on how well a user scores on different parts of the curriculum, the system would recommend a further study of particular parts of the knowledge network, thus acting as a kind of tutor to a student. When a question does not have a satisfactory answer within the knowledge base or the user disagrees with the answer that is offered, the user will be stimulated to mark out and—ideally—resolve this knowledge gap. This connects the teaching and assessment module to the research module: depending on how easy or ambiguous the question is, answering it can be more a questions of assessment of existing knowledge or of stimulating the creation of new knowledge.

Like Wikipedia, such a knowledge network should be able to grow indefinitely, so as to encompass all the relevant knowledge about a particular domain. Moreover, the system would automate the social and organizational functions of a university by stimulating users to develop their knowledge, assessing their progress, giving them feedback about how well they are doing, proposing concrete goals for their next study or research, and thus mobilizing and coordinating their cognitive efforts as efficiently as possible towards an ever increasing understanding of the domain.

If such a system were deployed at a global level, like Wikipedia, it would provide an invaluable aid to the cognitive development of the world population. Moreover, it would provide the ideal tool to further elaborate our transdisciplinary worldview, by helping us to cope with the complexity and fragmentation of existing, disciplinary knowledge.

Conclusion

After more than two decades of development, the conceptual framework underlying the ECCO approach has become very rich, coherent and comprehensive, extending from the most abstract, ontological foundations to the most concrete, practical applications. The worldview has been inherited to an important extent from the Principia Cybernetica Project [Heylighen, Joslyn & Turchin, 1991; Heylighen, 2000], with its evolutionary-cybernetic philosophy pioneered by Valentin Turchin [1977, 1993], and inspired by other great evolutionary, cybernetics and systems thinkers, such as Campbell [1974], von Bertalanffy [1973], Bateson [2000], and Ashby [1964]. This classic cybernetic tradition has been extended with the more recent approach of complex adaptive systems [Holland, 1996; Axelrod & Cohen, 1999], with its emphasis on the self-organization of multi-agent systems, and its applications to social, economic and ICT systems. The synthesis of both perspectives has led to a unique action-based ontology, where the elementary constituents of reality are conceived as actions and their agents, while the more complex phenomena are seen as self-organizing networks of such interacting agents, which gradually evolve to become ever more complex, adaptive and intelligent.

A first advantage of this ontology is that it steers clear of the perpetual philosophical controversies that arise from the mind-matter duality. Since the action ontology does not see reality as composed out of passive particles of matter whose behavior is fully determined by the laws of nature, it has no problem explaining how such particles can exhibit apparently intentional activity. It neither has any fundamental problem explaining the complex and seemingly mysterious abilities of the human mind, such as cognition, consciousness, and feeling, since it sees these as merely more evolutionarily advanced versions of the basic aspects of sensation, choice and goal-directedness that characterize even the simplest agents.

A second advantage of the "action first" approach is that it naturally leads to practical applications in the design of systems that act more effectively. One approach is to use our knowledge of self-organization in order to build systems composed of hardware or software agents that can function autonomously, adapting to the situation, however complex and unpredictable it may be. Another approach is to design what we called *mobilization systems*. These support the actions of human agents, helping them focus on the most important goals, while facilitating social self-organization. A final approach, the *stigmergic university* project, applies our philosophy to research and teaching, i.e. to the development and dissemination of knowledge across the boundaries of the traditional disciplines. The grand ambition here is to create a global conceptual network that is constantly being added to and refined by researchers around the world, while stimulating and guiding learners to assimilate the for them most interesting knowledge. This can be seen as a continuation of the original aims of the Principia Cybernetica Project [Heylighen, 2000], but now complemented by our novel insights in social and cognitive self-organization and in the psychology of human experience and emotion.

These applications are still mostly in the conceptual stage, with just a few prototypes available here and there as "proofs of concept". Obviously, ambitions as broad as these need a lot of time and effort to have any chance of being realized. Still, the example of Wikipedia—which evolved in less than a decade from a nice idea to the greatest knowledge network that ever existed and that is being used daily by millions of people—shows that a good design can surprisingly quickly turn into a global reality. Whether the present projects will be similarly successful is doubtful, to say the least, since success depends on so many other social, economical, technological and purely accidental factors than just having a good idea.

However, I believe that the fundamental value of our ideas will eventually become broadly recognized, thanks to the simplicity, comprehensiveness and coherence of our conceptual framework. Of course, this conceptual framework too needs further elaboration, and many gaps, ambiguities and perhaps inconsistencies still need to be resolved. In particular, while the applications are concrete, many concepts are still formulated at such a high level of abstraction that they may appear counterintuitive or difficult to grasp to those not acquainted with this way of thinking. This problem is gradually being resolved by adding more examples and intermediate concepts, bridging the gap between the most abstract, non-intuitive ideas and their concrete, real-world implementations. For example, the abstract notion of "fitness" (success in the most general evolutionary sense) can be connected to the concrete experience of "pleasure" via the intermediate concepts of "utility" and "happiness".

Yet, apart from the by now outdated Newtonian worldview, I do not know of any other philosophical system that offers such a broad, transdisciplinary scope of explanation—from elementary particles via subjective experience to the evolution of society—while being as simple, concrete, and practical in its application. I have tried to show how the system proposes inspiring and satisfactory answers to all the fundamental philosophical questions that define an *integrating worldview*, as formulated by [Aerts et al., 2002]. The appendix to this paper is a first attempt to systematically define and present the most important—of course, not all!—elements of this conceptual framework, so that the reader may be able to judge for herself whether this framework is as simple, coherent and comprehensive as claimed…

Appendix: Fundamental Concepts of the ECCO Worldview.

The following elaboration of the core concepts of our action ontology is structured like a glossary, where each new concept is followed by a definition with some further explanation. The definition normally only uses concepts that have already been defined before.

Agents and Evolution

action: a change in the state of affairs, from an initial state or "cause" to a subsequent state or "effect". This causal relation can be represented as a condition-action rule: whenever a certain condition (state functioning as *cause*) X is encountered, a particular action (change of that state, *effect*) is executed, deterministically or probabilistically, modifying that state into a new state Y. This can be represented simply as:

$$X \rightarrow Y.$$

Examples of very simple actions are chemical or physical reactions between particles or molecules, such as: $n \rightarrow p + e + \nu_e$ (beta decay in which a neutron n is converted into a proton, electron and neutrino), or $2H_2 + O_2 \rightarrow 2H_2O$ (production of water by the combination of hydrogen H with oxygen O).

Note that an action is intrinsically *relational*: it connects or links two states of affairs, and cannot exist independently of the phenomena it links. Also note that an action implicitly contains an arrow of time: there is a given direction, from cause or "before", to effect or "after". In some cases, the arrow may also point in the other direction, meaning that the change can be reversed, but such reversibility should not be assumed to be automatic. For example, the inverse reaction $2H_2O \rightarrow 2H_2 + O_2$ (electrolysis of water) is also possible, but requires some more stringent additional conditions (addition of energy) than the original reaction, so that it is a priori less likely to occur. While time is thus implicitly present in this ontology of action, I have proposed a method to derive it explicitly via a mathematical analysis of the network of actions [Heylighen, 1990, 2010], but this derivation is still incompletely worked out.

Although this definition explains action as a change of state, action is considered as the true primitive of our ontology, and therefore all other concepts, including "state" must be defined in terms of action or derived concepts.

state: The state of the world at a particular instant is defined by *the set of all actions that could be performed at that moment* [Turchin, 1993]. If actions have a certain probability to occur, like in quantum mechanics, then the state (like the wave function in quantum mechanics) includes the probability distribution of all these actions.

For example, the state n (representing a neutron) in the above example is defined in part by the fact that this state can transform into the new state $p + e + \nu_e$. Neutrons, as a type of particle, can participate in a limited number of particle reactions. The whole of these reactions determines the specific properties that define a neutron. Other particles, such as protons, participate in different reactions: the same actions cannot be applied to a state including a proton and to a similar state including a neutron.

For example, neutrons, having no electromagnetic charge, cannot participate in electromagnetic interactions, while protons can. The property of "charge", like all properties, is merely shorthand for the category of entities that have the potential of participating in a certain type of interactions (in this case electromagnetic interactions).

While this definition of state may seem counterintuitive, it can be mapped onto the more traditional notion of state as the set of all properties that are actual or "true" at a certain moment [Heylighen, 1990b], if we remember that a property needs to be observed to be deemed "true", and that an observation, as shown by quantum mechanics, is an action (which typically changes the observed state).

States are distinguished from each other by the fact that they allow different actions. On the other hand, actions are distinguished by the fact that they connect different states. Individual states and actions therefore distinguish or define each other, in a *bootstrapping* fashion. This principle can be expressed formally by what I have called the Bootstrapping Axiom [Heylighen, 1990, 2001]. The more general notions of state and action too are here defined in a bootstrapping manner: an action is a change of state; a state is a collection of potential actions. Whenever one of these potential actions is executed, the state by definition changes, since this particular action is no longer potential, while new actions may now have become potential.

agent: a persistent producer of actions. An agent is a condition, i.e. a part or aspect of a state. This condition is *persistent* in the sense that it conserves its identity (remain distinguishable) across subsequent states in a series of actions; it is a *producer* in the sense that it is necessary for these actions to occur. An example in chemistry is a catalyst, i.e. a molecule that participates in a reaction while remaining itself unchanged. The functioning of an agent can be expressed by an action of the following form.

$$A + X \rightarrow A + Y.$$

Here A is the agent or catalyst that is necessary for the reaction to occur, but which is not itself acted upon. In practice, this representation can be simplified to: agent A executes the condition-action rule $X \rightarrow Y$, i.e. whenever A encounters the condition X, it reacts in such a way that the new condition Y is produced.

Agents are the stable components of reality. Typical agents are people, animals, robots, organizations or cells, but agents can also be much simpler objects, particles or molecules. We saw that actions are in general directed towards one outcome rather than another one.

Therefore, agents have implicit preferences for certain actions over others, in the sense that when offered a “choice”, they are more likely to perform the "preferred" actions. Preference functions like what in physics is called a gradient or force field, pushing the agent in a particular direction.

event: an action that is not produced by an agent [cf. Turchin, 1993]. An example is the neutron decay we just discussed, or the mutation of a molecule of DNA. While an agent may be involved in that event (like an organism is typically involved in DNA mutation), the presence of that agent is not necessary for the action to take place. For example, a cosmic ray may damage a DNA molecule independently of whether that molecule is part of a living cell. While one may be tempted to see the cosmic ray (a high-energy photon) as the agent of that mutation because it produced that event, it is not persistent, and therefore not considered as an agent in the present framework: the photon is absorbed by the molecule and no longer exists after the event. Therefore, this event may have a *cause* (the encounter of a photon and a DNA molecule), but it does not have an *agent*. In conclusion, we may note that agents are not in control of events, only of the actions that they themselves initiate.

goal: one of the end states or *attractors* to which an agent's actions is likely to lead, i.e. a state where the agent’s preference is for not further changing that state. Goals in this sense are most of the time implicit: the agent does not have any representation or awareness of what that end state is. Yet, for an external observer it may quickly become apparent that all the actions of the agent are directed towards the same end state.

For a physical object, the goal is typically to minimize potential energy or free energy, i.e. maximizing equilibrium or stability. Therefore, the goals of physical systems are normally “potential wells”: the locally most stable states. For a living organism, the goal is to maximize biological fitness, i.e. survival and reproduction.

While it may seem strange to interpret an attractor or equilibrium as a goal, in practice goal-directed behavior cannot be distinguished from attractor-directed behavior: both eventually end up in the attractor or goal, and do so even when the agent is disturbed, i.e. pushed into a different state than the one it had acted towards. This behavior is called *equifinality* [von Bertalanffy, 1973]: different initial states lead to the same final state. We will typically attribute a goal or intention to an agent if we see that this agent is constantly working towards this goal, and continues to do so even if we try to prevent it from doing so by changing its course. Interpreting attractors as goals or desired states is a direct application of what we have called “the intentional stance”.

fitness: a property that measures the degree of recurrence of a state or agent: the higher the fitness, the more likely we are to encounter that same condition again. Fitness can be achieved by stability or survival (the state persists under further actions), or by multiplication (the state is recurrently produced in the right circumstances, either by replication of an existing template, or by independent recreation). Fitness is the ultimate measure of success of a state:

by definition, unfit states do not recur and therefore are eliminated from the scene. An agent can be seen as a state with a certain, minimal degree of fitness that allows it to persist at least during certain actions. Maximizing fitness is the implicit goal or preference of all agents, since the attractor of their actions by definition is a state that maximally recurs.

utility: a measure of the "success" of an action, i.e. the degree to which the action has made the agent increase its fitness or advance towards its goals. Intuitively, we can see it as the amount of "benefit" or "satisfaction" that an agent obtains from an action. When confronted with different options for actions, agents will normally choose the one from which the highest utility can be expected, i.e. the one that leads to the goal most directly. For physical systems, decrease of free energy is a good measure of utility. For living systems, enhancement of biological fitness is such a measure. For people, utility typically takes the form of the happiness or profit that an agent gets from an action.

cost: the amount of utility consumed or wasted by performing an action. Actions are typically subjected to trade-offs: improvement in one aspect tends to be counterbalanced by deterioration in another aspect. For example, an action normally uses energy or some other resource, and some of this energy will be wasted or dissipated and therefore no longer available to perform further actions. Even if the action was successful, the agent will have lost some of its power to perform subsequent actions, and in that sense the action had a cost.

variation: the on-going change in the state of world caused by subsequent actions. Variation produced by events is in general random, which means that it is difficult to a priori discern the direction towards which it moves. Variation produced by an agent is normally directed towards the agent's goal. However, since the agent has at most a short-term, local knowledge of the effects of its actions (bounded rationality, see further), there is no guarantee that variation will reach the goal in the long term. Therefore, variation is always to some degree *blind*, since no agent can foresee all the consequences of its actions.

natural selection: the selective retention of a particular state, because no further actions occur that change that state. Such a state corresponds to what we have called an attractor of the dynamics, or a state with high fitness. Practically any non-trivial dynamics will have several attractors. Therefore, if you wait long enough, variation will eventually reach such an attractor, implying that the state it produces is selected from among the other, non-attractor states, which are thereby eliminated. When the dynamics is dominated by agents, such a state typically corresponds to a (local) maximum of the agents’ fitness function.

evolution: the long-term, directed change in the state of world towards higher overall fitness which results from the interplay of the (partly random, partly directed) variation of states and the eventual selection of states with higher fitness. Evolution can be seen as a search for fitness based on trial-and-error, where variation produces the trials, and selection eliminates the errors [Heylighen, 1992, 1999a, 2007c].

challenge: any difference between the agent's present situation and its preferred situation or goal [called "problem" in Heylighen, 1990b]. To resolve that difference, the agents must decide about and perform a sequence of actions. Therefore, a challenge is the main motivator or stimulator of action for an agent. In practice, an agent will be confronted with challenges most of the time, as its present situation can almost always be better in some way, thus inviting the agent to attempt to achieve that improvement by choosing the most appropriate actions.

course of action: the sequence of actions that an agent would perform if left undisturbed, and the resulting trajectory through its state space that the agent would follow. The course of action will normally increase fitness, i.e. bring the present state closer to a goal state, and this as directly as possible. The course of action can be seen as an intended or anticipated sequence of actions. In practice, however, what is intended and what really happens tend to diverge, as we will now analyze in more detail.

diversion: any change in the agent's situation that makes the agent deviate from its present course of action. This deviation can be positive (moving it closer to the goals), negative (moving away from the goals), or neutral. The defining characteristic of a diversion is that the agent has no control over it (although the agent may try to control its subsequent effects): it does not originate in the agent's decision-making, but is unexpected, coming from an initially unforeseen origin. A diversion normally creates a new challenge: how to deal with this unforeseen situation? Examples are a sudden discovery, an obstacle appearing on the road, an apple falling from a tree, an unexpected phone call, a fluctuation.

disturbance: a negative diversion; i.e. a phenomenon that, if left unchecked, would make the agent's state deviate further from its goals, and thus reduce its fitness. Disturbances typically originate in the environment, but can also appear because of some malfunctioning within the agent itself. Examples are obstacles, errors, accidents, encounters with predators, parasites or otherwise hostile agents, diseases, poor weather conditions, etc. Serious disturbances need to be dealt with; otherwise they may endanger the agent's survival.

affordance: a positive diversion [Heylighen & Vidal, 2008]. An initially unforeseen change in the situation that creates an opportunity for the agent to increase its utility, so that it can reach its goals more quickly or easily than expected. Affordances can be the discovery of tools, means or resources (e.g. a phone, a hammer, food, someone that can give advice) that help the agent achieve its goals. They can also be realized negatively, as the disappearance of expected obstacles or constraints (e.g. a clearing up of the weather, a reduction in the price of energy).

counteraction: an action performed by an agent that suppresses or compensates for a disturbance, so as to reduce the deviation from the goal or course of action that the disturbance created.

regulation (or **control)**: the process by which an agent continually neutralizes deviations from its goals, by effectively counteracting disturbances [Heylighen & Joslyn, 2001]. Regulation implements negative feedback: deviations in one direction are compensated by reactions that push the state in the opposite direction, so as to reduce their effect. The classic example of regulation is the functioning of a thermostat, which switches on the heating as soon as the temperature moves below its set temperature or goal, and switches off the heating as soon as the temperature moves above the goal temperature.

exploitation: the use of known affordances in order to maximize the increase in utility they can bring about. This requires that the agent perform the right actions to extract utility from the affordance. Examples are collecting fruit, mining for coal, driving a car, cultivating crops, or asking a passerby for directions.

exploration: the process by which an agent searches for as yet unknown affordances, by trying out actions without specific expectation of their results, in the hope that one of them would uncover an affordance. Examples are animals foraging for food, children playing, scientists designing and performing experiments, readers browsing a magazine or surfing the Internet, or explorers mapping out new territories.

the exploration-exploitation trade-off: the difficult decision for an agent about how much energy to invest in exploration rather than exploitation. While exploitation of known affordances makes the agent advance to the goal most reliably, affordances can become exhausted, lose their usefulness because of a change in the situation, or lose their competitive edge relative to new affordances. Therefore, it is wise to invest in discovering new affordances before the old ones have lost their power. But exploration alone is a too risky and inefficient strategy, and must be complemented by exploitation. While there does not seem to be an optimal method to resolve the exploration-exploitation trade-off, a general rule seems to be that a more variable, unpredictable environment will necessitate more exploration, since new affordances are likely to appear while known ones may disappear. A more stable environment with known, high-quality resources lends itself more to exploitation.

navigation: the process by which an agent constantly adjusts its course of action so as to dynamically maximize its advance in utility while taking into account the diversions it encounter. This means that the agent should optimally allocate its effort in simultaneously counteracting disturbances (regulation), finding affordances (exploration), and making use of the affordances it has found (exploitation). Navigating means that the agent needs to set out a well-thought out, but flexible course, using a combination of planning (to deal with known disturbances and affordances) and improvisation (to deal with new diversions).

Cognition

variety of action: the number of actions that an agent can potentially execute. The larger an agent's variety of action, the larger the variety of diversions that the agent can deal with, because different types of diversions typically require different types of actions. This is a generalization of Ashby’s [1958, 1964] well-known “law of requisite variety” [Heylighen, 1992]. However, while increasing the variety of actions makes the agent potentially more powerful in achieving its goals, it also makes it more difficult for the agent to select the most appropriate action.

uncertainty: the degree to which an agent is unsure about what to do or to expect. The larger the number of options that can potentially occur, the larger the uncertainty, and therefore the larger the amount of trial-and-error that the agent will have to perform before it is likely to achieve a high-utility decision. Agents with high uncertainty will therefore be very inefficient in accumulating utility. Uncertainty is traditionally measured using Shannon's formula for entropy, which is based on the probability distribution of the different options [Heylighen & Joslyn, 2001].

information: any change of condition, functioning as a signal, that reduces uncertainty for an agent. Using Shannon's formula, the amount of information in a message can be calculated as the initial uncertainty minus the new uncertainty (after the message has been received). Note that information is always relative to an agent and the state or situation in which that agent is: what may be informative in one case, is likely to be meaningless, irrelevant or redundant in another case. A signal does not contain information on its own: some agent needs to process or interpret that signal in a way that allows the agent to reduce its uncertainty. Therefore, information should not be interpreted as a *substance* similar to mass or energy.

intelligence: given a certain information, the degree to which an agent is able to make good decisions, i.e. selections of actions that maximally accumulate utility in the long term. A zero-intelligence agent is one that selects actions at random from the ones allowed by the remaining uncertainty. Intelligence has two components: knowledge (or "crystallized intelligence"), and fluid intelligence.

knowledge: the ability, typically derived from experience, education or communication, to anticipate the consequences of a given state or action. Knowledge can be represented in the form of condition → action or condition→condition rules. The former specifies which action the agent should perform in a given condition, the latter specifies which new condition can be expected to follow this given condition. Knowledge differs from information in that it produces general predictions or expectancies, applicable in many different situations, while information strictly speaking only applies to the present situation.

fluid intelligence: the ability to internally explore many different combinations of possible events and actions in order to find the one that according to the existing knowledge would produce the largest utility. This requires at least a mechanism of inference, such as the concatenation of condition-condition rules, e.g. A → B, and B→ C, therefore A → C. By thus inferring the expected effects from different potential conditions and actions, the agent selects the ones that seem most appropriate, in this way planning or setting out a course of action that may extend well into the future, and that takes into account a variety of alternative scenarios.

cognition: the acquisition, processing, storage, and use of information and knowledge to support intelligent decision-making [Heylighen, 2007d]. This includes perception (the processing and interpretation of incoming information), learning (extracting recurrent regularities from perceptions and storing them into memory in the form of knowledge), and inference (using stored knowledge patterns to anticipate situations as yet not perceived).

rationality: the hypothetical desire and ability of an agent to always choose the best action. In reality, rationality is restricted or bounded, as an agent never has enough information, knowledge or intelligence to accurately determine the utility of all possible courses of action. Bounded rationality implies that there is always an element of uncertainty or trial-and-error involved in making decisions; no decision can be a priori proven to be the best one.

intelligence amplification: a process that increases the ability of an agent to make good decisions. Intelligence can be amplified by providing more or better knowledge (e.g. an encyclopedia in which facts can be checked), by increasing the ability to explore many different possibilities (e.g. by means of a computer program that can make more and faster inferences than a human brain, or via drugs that improve thinking in the brain), or by some combination of these.

Interactions

interaction: reciprocal effect of two (or more) agents (say, A and B) on each other: the action performed by A creates a condition that triggers another action (reaction) from B. This second action in turn affects the condition of A, stimulating it to react in turn, and so on. Interaction can go on indefinitely, or stop when the final condition does not trigger any further action.

zero-sum interaction: an interaction in which every gain in utility for one agent is counterbalanced by an equal loss in utility for the other agent [Wright, 2001]. This typically occurs when utility is proportional to the amount of "material" resources (such as food, money or energy) that an agent acquires: when the total amount of resources is conserved, the sum of gains (positive changes) and losses (negative changes) must equal zero. An example without material resources is a chess game or a sports competition: if one agent wins, then by definition the other agent(s) lose(s). The exchange of resources, such as information, that do not obey a conservation law will in general not lead to a zero-sum interaction: the sum can be positive (overall gain) or negative (overall loss).

synergy: gain in overall utility caused by an interaction. Synergy is characteristic of an interaction with a positive sum [Wright, 2001; Corning, 1998]. A primary example is a "win-win" situation where all parties gain in utility. This typically happens when the action performed by one agent to advance towards its goals makes it easier for another agent to achieve its goals as well, by removing a shared disturbance or making available a common affordance. An example is the sharing of information or knowledge so that a discovery made by one agent can benefit the other agents as well. Unlike matter or energy, information is not conserved, and therefore an information gain for one agent can be accompanied by a similar gain for the other.

friction: a loss in overall utility caused by interaction [Heylighen, 2007c; Gershenson, 2007]. This is the opposite of synergy, and characteristic of an interaction with negative sum, where all parties together lose (although one may gain at the expense of a larger loss by the others). This typically occurs when resources are dissipated or wasted during the interaction. An example is a traffic jam, where enormous amounts of fuel, time and energy are wasted because of mutual obstruction between vehicles. The dissipation can be physical (dissipation of energy or thermodynamic entropy, because of diffusion or physical friction), or informational (waste of resources because of uncertainty leading to many trials ending in error).

cooperation: the relation between agents involved in a synergetic or positive sum interaction [Heylighen & Campbell, 1995; Heylighen, 2004]. Usually, cooperation is assumed to be intentional, i.e. the agents act in the expectation of a positive sum result (now or later). For example, wolves may cooperate in hunting a prey animal (such as moose) that is too large to be killed by any wolf individually, but small enough to be killed by a pack of wolves. If the positive sum interaction is unintentional, we may just call it "synergy".

competition: the relation between agents involved in a zero-sum interaction [Heylighen & Campbell, 1995; Heylighen, 2004]. Since the total amount of resources to be gained is fixed, a better result for one agent necessarily implies a worse result for the other agent(s). Thus, competition tends to exacerbate the differences in success between agents, and therefore the natural selection that lets the fitter agents proliferate while eliminating the less fit.

conflict: the relation between agents involved in an interaction with friction or negative sum [Heylighen & Campbell, 1995; Heylighen, 2004]. This happens when the goals of the different agents are inconsistent: movement towards the one implies distancing towards the other one. Usually, conflict is assumed to be intentional, i.e. the agents act in the expectation of inflicting a loss on the other party. If a negative sum interaction is unintentional, like in a traffic jam, we may just call it "friction".

transaction costs (or **interaction costs)**: the degree to which utility in a positive-sum interaction is lost to friction. Even when the interaction overall is synergetic, some of the generated utility will be dissipated during the process. Typical transactions costs are the effort invested in finding the right partner to interact with, negotiating who will contribute what to the transaction, and making sure that everything happens as planned [Martens, 2004]. According to some estimates, in our present economic system more than half of economic value generated is lost to transaction costs. A fundamental source of transaction costs is uncertainty: since the agent does not know what transaction to enter into, what to agree upon, or what to expect, it will need to spend a lot of energy in search, negotiation, and enforcement of agreements.

Systems

constraint: a limitation of the variety of action for some agent(s). A constraint makes certain actions impossible, and thus reduces the "freedom" of the agent. A constraint also makes the further sequence of actions more predictable by reducing uncertainty. Putting someone in prison imposes a constraint on that person, but so does enunciating a law or contract that that person has to obey. Attaching a ball to a chain is another example of a constraint, this time on the ball. An attractor is a spontaneously evolved constraint: once an agent has entered an attractor, it loses the freedom to visit states outside that attractor.

bond or **connection:** a stabilized interaction between two agents exerting a constraint on their further action. A bond is an attractor for the dynamics of interaction, and thus the result of variation and selection. Examples are chemical bonds holding together atoms in a molecule, the biological bond between cells in a multicellular organism, and a marriage, as a bond between two people. Such bonds reduce the relative freedom for the agents: they cannot perform certain actions without the other agent "going along" with that action. For example, an atom in a molecule cannot move further than a certain distance from the other atoms. In order for it to move more, the other atoms need to move along with it.

system: a group of agents held together by a shared constraint or network of bonds. This cohesion distinguishes it from the environment, which groups any other agents with which there is a weak(er) interaction. The agents in the system can be seen as the system's components. Insofar that the agents in the system share a goal, the system functions like a higher-order agent. A multicellular organism is an example of such a superagent consisting of cells as individual agents.

supersystem: a system whose components are systems themselves. These components are called the "subsystems" of the original system. In general, any system contains subsystems, and is contained in one or more supersystems. Thus, it forms part of a chain or hierarchy, extending upwards towards ever larger, more encompassing wholes, and downwards towards ever smaller, more primitive parts. For example a human being is part of a city, which is part of a country, which is part of global society, while having as parts organs, which have as parts cells, which have as parts molecules, etc.

supersystem transition: the emergence of a supersystem by the creation of bonds between existing systems. This can be seen as a discrete evolutionary step towards a higher level of complexity.

environment: everything that is considered to be external to a given agent or system, but that still interacts with it. This means that the environment provides the initial conditions or input that trigger an interaction, while accepting the output or change produced by that action.

complex adaptive system: a system consisting of many interacting agents, where their interactions are not rigidly fixed, preprogrammed or controlled, but continuously adapt to changes in the system and in its environment. Examples are ecosystems, communities and markets. The bonds between such agents are relatively weak and flexible, so that there is still a lot of freedom for the system to adapt. On the other hand, the agents do depend on each other, and therefore their individual freedom is limited. For example, a firm in a market cannot afford to set prices much higher than those of the other firms without going bankrupt, but this constraint is not explicitly fixed.

collective intelligence: the degree to which the agents in a system collectively can make good decisions as to their future course of action; in particular, the degree to which the agents collectively can make better decisions than any of them individually [Heylighen, 1999b]. Collective intelligence typically arises from the pooling of knowledge: since each agent's knowledge is slightly different from the knowledge of the others, all agents together will have access to a larger pool of more diverse knowledge. Therefore, collective intelligence can be amplified by increasing the diversity of knowledge among the agents.

distributed cognition: the acquisition, storage and use of information and knowledge distributed over different agents in a system, so as to support their collective intelligence [Heylighen, Heath & Van Overwalle, 2004]. This requires mechanisms to communicate information across the different agents in a coordinated manner, so that the right information is gathered at the right place at the right time. Examples of such supporting mechanisms for the exchange of information are pheromones (smell signals), language, books, signs, computers and the Internet.

medium: the substrate that carries or supports the interactions between agents; that part of the world that is changed by an agent's action, and whose changed state is perceived as a condition for a subsequent action by another agent. Examples of media are air for acoustic interaction, the electromagnetic field for electric interactions, the physical surroundings for collaborative building, a wiki website for collaborative writing. The medium is often the environment shared by the interacting agents, but can also be internal to the agents.

Organization

coordination: the arrangement or mutual alignment of actions so as to maximize synergy and minimize friction in their overall pattern of activity. It implies that any two actions performed simultaneously or subsequently are selected so as to maximally complement and minimally obstruct each other. This requires a minimization of the uncertainty that otherwise would dissipate resources in needless trial-and-error, and therefore the imposition of appropriate constraints or bonds that drive the action in the right direction.

There are two types of basic relational constraints between actions: **parallel**, specifying which actions could or should go on simultaneously, and **sequential**, specifying which action could or should follow which other action. An example of sequential coordination is planting a tree: first a hole must be dug, then the tree must be inserted, then the remaining hole must be filled with earth. There is no way to achieve the desired result by changing the sequence of these actions, because otherwise the one action will obstruct the other one. An example of parallel coordination is two or more people pushing simultaneously in the same direction to move a heavy object: if they would have pushed in different directions, or at different times, the object might not have moved, or the movement would at least have suffered much more friction.

self-organization: the spontaneous emergence or evolution of coordination in a complex adaptive system [Heylighen, 2002]. Self-organization reduces variety or uncertainty, and thus imposes constraint. The driving force behind self-organization is the co-evolution or mutual adaptation between the different agents in the system: actions and reactions produce a continuously changing configuration of interactions (variation); however, the more synergetic a configuration, the more "satisfied" the agents will be with the situation, and thus the less they will act to produce further changes (selective retention or preference for synergetic configurations); vice versa, the more friction there is, the more the agents will be pressured to intervene and change course in order to increase their utility (elimination of high friction configurations). Thus, self-organization is merely an application of the evolutionary dynamic of variation (because of actions triggering further actions) and natural selection (because of the implicit preference of agents for the more synergetic patterns of action).

organization: a stabilized network of interactions between agents that functions to ensure the coordination of their actions. This structure specifies the specific roles of and interactions between the system's agents. Its function is to maximize synergy and minimize friction (including transaction costs) in their further interactions. For example, in a human organization the different individuals each have their own responsibilities, and the rules of the organization specify who interacts with whom in what way. This minimizes transaction costs, since it is no longer necessary to search for partners, negotiate with them, or strictly monitor whether they do what they are expected to do. An organization can be imposed from the outside (like in a system engineered by a designer or controlled by a manager), or emerge from self-organization.

mediator: a regulatory structure external to the agents that promotes coordination between them [Heylighen, 2004; Gershenson, 2007]. An example is the system of roads, traffic lights, traffic signs, and lanes that coordinates the movement of vehicles so as to minimize mutual obstruction (i.e. friction). Mediation may emerge from self-organization (e.g. vehicles spontaneously moving to the side in order to let others pass), or be imposed by an inside or outside agent (e.g. a policeman regulating traffic).

stigmergy: a form of indirect coordination via the medium, where the trace left by an action in the medium stimulates the performance of a subsequent action, thus building further on the work that has already been done [Heylighen, 2007ab; Parunak, 2006]. Examples of stigmergic interaction are termites collectively building a termite hill by adding mud to heaps left by other termites, ants mapping out their territory by leaving trails of pheromone that lead other ants to food sources, and people adding their insights to the writings of others, thus collaboratively developing an encyclopedia (Wikipedia) on the web.

Stigmergy is typically the result of the self-organization of a mediator out of the medium. It is probably the simplest way to achieve coordination in a complex system because it does not make any cognitive demands on the agents (such as remembering who is to do what when), and therefore functions even with agents of very low intelligence, such as insects. It also does not require any relationship, communication or even direct interaction between the agents: contributors to a Wikipedia page typically have no idea who the other contributors are, yet their contributions are seamlessly integrated into the emerging whole.

evolution of cooperation: the general tendency for interactions to become more synergetic through variation and selection, thus reducing competition and conflict [Heylighen & Campbell, 1995; Stewart, 2000]. The biggest obstacle to the evolution of cooperation is the problem of "free riders": agents that profit from the collaborative effort of other agents but without contributing themselves. Since free riders typically extract more benefit from the cooperation than the cooperators themselves, they risk outcompeting the cooperators, thus destroying any tendency to cooperation. This problem can be overcome via the evolution of a mediator that makes free riding impossible, or at least less attractive than true cooperation.

References

Aerts, D. et al. (2002). Worldviews: from fragmentation to integration. Brussels: VUB Press, Online access http://www.vub.ac.be/CLEA/pub/books/worldviews.html.

Ashby, W. R. (1958). Requisite variety and its implications for the control of complex systems. Cybernetica, 1(2), 83-99.

Ashby, W.R. (1964). An introduction to cybernetics, Methuen London.

Axelrod, R., and M. D. Cohen. (1999). *Harnessing Complexity: Organizational Implications of a Scientific Frontier*. Free Press.

Bateson, G. (2000). *Steps to an Ecology of Mind*. University of Chicago Press.

Battram A. (1996). Navigating Complexity: The Essential Guide to Complexity Theory in Business and Management.

Bernheim, J. (1999). The Cognitive Revolution and 21st Century Enlightenment: towards a contemporary evolutionary progressive world view. in: Science, Technology, and Social Change: The Orange Book of" Einstein Meets Magritte", 63.

Bonsack, F. (1977). Invariance as a Criterion of Reality. *Dialectica* 31, no. 3: 313–331.

Boulding, K.E. (1956). General Systems Theory-The Skeleton of Science. Management Science, 2(3), 197-208.

Brier, S. (2008). Cybersemiotics: why information is not enough!, Univ. of Toronto Press.

Campbell D.T. (1974). "Evolutionary Epistemology", in: The Philosophy of Karl Popper, Schilpp P.A. (ed.). (Open Court Publish., La Salle, Ill.), p. 413-463.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2, no. 3: 200–219.

Charlton, B. (2002). What is the Meaning of Life? Animism, Generalised Anthropomorphism and Social Intelligence. http://www.hedweb.com/bgcharlton/meaning-of-life.html.

Charlton, B. (2007). Alienation, recovered animism and altered states of consciousness. *Medical Hypotheses* 68, no. 4: 727-731. doi:10.1016/j.mehy.2006.11.004.

Combs, A., and S. Esbjörn-Hargens. (2006). An Integral Tour of Consciousness Studies. *Journal of Integral Theory and Practice, V1. 1*: 162.

Cools S., Gershenson C., D'Hooghe B. (2007). Self-organizing traffic lights: A realistic simulation, In Prokopenko, M. (Ed.). Self-Organization: Applied Multi-Agent Systems, Chapter 3, pp. 41-49. Springer, London.

Corning, P.A. (1998). “The synergism hypothesis”: On the concept of synergy and its role in the evolution of complex systems. Journal of Social and Evolutionary Systems, 21(2), 133–172.

Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. Harper & Row.

Dennett, D. C. (1989). *The intentional stance*. The MIT Press.

Diener E. & R. Biswas-Diener (2008): Happiness: Unlocking the Mysteries of Psychological Wealth, Wiley-Blackwell.

Elmenreich, W., R. D'Souza, C. Bettstetter, and H. de Meer. (2009). A Survey of Models and Design Methods for Self-Organizing Networked Systems. *Self-Organizing Systems: 4th IFIP TC 6 International Workshop, IWSOS 2009, Zurich, Switzerland, December 9-11, 2009, Proceedings*: 37–49.

Gendlin, E. T. (1962). *Experiencing and the creation of meaning*. Free Press of Glencoe New York.

Gershenson C. & F. Heylighen (2004). How can we think the complex? in: Richardson, Kurt (ed.) Managing the Complex Vol. 1: Philosophy, Theory and Application.(Institute for the Study of Coherence and Emergence/Information Age Publishing)

Gershenson C. (2007). Design and Control of Self-Organizing Systems (ECCO working paper 2007-01, also PhD thesis, Faculty of Sciences, Vrije Universiteit Brussel)

Gershenson C., Heylighen, F. (2004): Protocol Requirements for Self-organizing Artifacts: Towards an Ambient Intelligence, in: Proc. Int. Conf. on Complex Systems (New England Institute of Complex Systems)

Geyer, F. (1994). Alienation, participation and increasing societal complexity. Kybernetes 23: 10- 34.

Hassenzahl, M., and N. Tractinsky. (2006). User experience–a research agenda. *Behaviour & Information Technology* 25, no. 2: 91–97.

Hawkins, J., and S. Blakeslee. (2005). *On intelligence*. Henry Holt and Company.

Heath, C. & Heath, D. (2007). Made to Stick: Why Some Ideas Survive and Others Die, Random House.

Heath, J. & Anderson, J. (2010), Procrastination and the Extended Will, in: Andreou, C. & White, M. (eds.) The Thief of Time: Philosophical Essays on Procrastination, Oxford Univ Press.

Heylighen F. & Bernheim J. (2000): "Global Progress II: evolutionary mechanisms and their side-effects", Journal of Happiness Studies 1(3), p. 351-374

Heylighen, F. & Campbell, D.T. (1995). Selection of Organization at the Social Level: obstacles and facilitators of metasystem transitions. World Futures: the journal of general evolution, 45, 181–212.

Heylighen, F. & Gershenson, C. (2003). The Meaning of Self-organization in Computing. IEEE Intelligent Systems 18:4, p. 72-75.

Heylighen, F. & Joslyn, C. (2001). Cybernetics and Second Order Cybernetics. Encyclopedia of Physical Science & Technology, 4, 155-170.

Heylighen, F. & Vidal C. (2008). Getting Things Done: The Science behind Stress-Free Productivity. Long Range Planning 41, pp. 585 – 605.

Heylighen, F. (1989). Causality as Distinction Conservation. A Theory of Predictability, Reversibility, and Time Order. *Cybernetics and Systems* 20, no. 5: 361-384.

Heylighen, F. (1990). A structural language for the foundations of physics. International Journal of General Systems, 18(2), 93–111.

Heylighen, F. (1990)b. Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science, (Communication & Cognition, Gent),

Heylighen, F. (1992). Principles of Systems and Cybernetics: an evolutionary perspective. Cybernetics and Systems' 92, 3-10.

Heylighen, F. (1997). Objective, subjective and intersubjective selectors of knowledge. *Evolution and Cognition* 3, no. 1: 63-67.

Heylighen, F. (1999)a. The Growth Of Structural And Functional Complexity During Evolution. in: Heylighen, Bollen & Riegler (eds.) The Evolution of Complexity: The Violet Book of" Einstein Meets Magritte (Kluwer).

Heylighen, F. (1999)b: "Collective Intelligence and its Implementation on the Web: algorithms to develop a collective mental map", Computational and Mathematical Organization Theory 5(3), 253-280.

Heylighen, F. (2000): "Foundations and Methodology for an Evolutionary World View: a review of the Principia Cybernetica Project", Foundations of Science, 5, p. 457-490.

Heylighen, F. (2001): "Bootstrapping knowledge representations: from entailment meshes via semantic nets to learning webs", Kybernetes 30 (5/6), p. 691-722.

Heylighen, F. (2002). The Science of Self-organization and Adaptivity. The Encyclopedia of Life Support Systems.

Heylighen, F. (2004): "Mediator Evolution: a general scenario for the origin of dynamical hierarchies",

Heylighen, F. (2007)a. Warum ist Open-Access Entwicklung so Erfolgreich? (Why is Open Source Development so Successful? Stigmergic organization and the economics of information, in: B. Lutterbeck, M. Baerwolff & R. A. Gehring eds., Open Source Jahrbuch 2007, Lehmanns Media, 2007, p. 165-180.

Heylighen, F. (2007)b: "The Global Superorganism: an evolutionary-cybernetic model of the emerging network society", Social Evolution & History. 6 No. 1, p. 58-119

Heylighen, F. (2007)c: Accelerating Socio-Technological Evolution: from ephemeralization and stigmergy to the global brain, in: "Globalization as an Evolutionary Process: Modeling Global Change", edited by George Modelski, Tessaleno Devezas, and William Thompson, London: Routledge, p. 286-335.

Heylighen, F. (2007)d Cognitive Systems: a cybernetic perspective on the new science of the mind (ECCO working paper 2007-07) , lecture notes for the course "Cognitieve Systemen"

Heylighen, F. (2008). Complexity and Self-organization, in: Encyclopedia of Library and Information Sciences, eds. M. J. Bates & M. N. Maack Taylor & Francis, 2008

Heylighen, F. (2009) Stimuleren van geluk en sociale vooruitgang: een libertair paternalistische benadering (ECCO Working Paper 2008-14), to appear in Ethiek en Maatschappij.

Heylighen, F. (2010) The Self-organization of Time and Causality: steps towards understanding the ultimate origin, Foundations of Science.

Heylighen, F. Heath M., F. Van Overwalle (2004): The Emergence of Distributed Cognition: a conceptual framework, Proceedings of Collective Intentionality IV, Siena (Italy)

Heylighen, F. Joslyn, C. & Turchin, V. (1991). A Short Introduction to the Principia Cybernetica Project. Journal of Ideas, 2(1), 26-29.

Heylighen, F. P. Cilliers, & C. Gershenson (2007): "Complexity and Philosophy", in: Jan Bogg and Robert Geyer (editors), Complexity, Science and Society. (Radcliffe Publishing, Oxford

Holland, J. H. (1996). Hidden order: How adaptation builds complexity. Addison Wesley Publishing Company.

Kharatmal, M. & Nagarjuna, G. (2009). SELF Platform: A Teacher-Centric Collaborative Authoring System. Journal of Applied Collaborative Systems Vol, 1(2).

Kharatmal, M. & Nagarjuna, G., Refined concept maps for science education: A feasibility study. Episteme, 3, 76–80.

Kiemen, M. (2006). A triple loop model to ground higher-level cognition. In *In Proceedings of the 18th European Meeting on Cybernetics and systems Research*, 2:369–374.

Kiemen, M., T. Coenen, E. Torfs, and E. Vandijck. (2009). Enterprise innovation planning with social software. In *In Proceedings of the 20th ISPIM conference-The future of innovation*.

Klein, M. (2007). Achieving Collective Intelligence via Large-Scale On-line Argumentation. In Internet and Web Applications and Services, 2007. ICIW'07. Second International Conference on, 58–58.

Lyubomirsky, S. (2008). The how of happiness. Penguin Press.

Lyubomirsky, S., L. King, and E. Diener. (2005). The benefits of frequent positive affect: does happiness lead to success? Psychological Bulletin 131: 803–855.

Malone, T. W, and M. Klein. (2007). Harnessing collective intelligence to address global climate change. Innovations: Technology, Governance, Globalization 2, no. 3: 15–26.

Martens B. (2005). The cognitive mechanics of economic development and institutional change. (Routledge).

Martin M.A. (2009). Teacherplex.com: A New Style of Website Devoted to Education (ECCO working paper 2009-06)

Maturana, H.R. & Varela, F.J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*, Shambhala.

Menghini, F., N. van Rijsbergen, and A. Treves. (2007). Modelling adaptation aftereffects in associative memory. *Neurocomputing* 70, no. 10 (June): 2000-2004. doi:10.1016/j.neucom.2006.10.081.

Nagarjuna, G. (2007) GNOWSYS: A system for semantic computing, 2005, available at: http://www.gnu.org/software/gnowsys/GNOWSYS_files/cpv2/index.html.

Nagel, T. (1979), Panpsychism, in *Mortal Questions* (Cambridge: Cambridge University Press).

Parunak, H. V. D. (2006). A survey of environments and mechanisms for human-human stigmergy. *Environments for Multi-Agent Systems II*: 163-186.

Prigogine, I. (1980). *From being to becoming. Time and complexity in the physical sciences.* New York: Freeman.

Rescher, N. (1996). *Process metaphysics: An introduction to process philosophy*. State Univ of New York Pr.

Rochat, P. (2001). Origins of self-concept. *Blackwell handbook of infant development* 2: 191–212.

Rodriguez M.A., Watkins J. (2009) Faith in the Algorithm, Part 2: Computational Eudaemonics, Proceedings of the International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Lecture Notes in Artificial Intelligence, Springer-Verlag.

Rodriguez, M.A. (2008), Grammar-Based Random Walkers in Semantic Networks, Knowledge-Based Systems, volume 21, issue 7, pages 727-739.

Sheldon, K. M, and S. Lyubomirsky. (2006). Achieving Sustainable Gains in Happiness: Change Your Actions, not Your Circumstances. Journal of Happiness Studies 7, no. 1: 55–86.

Solms, M., and O. Turnbull. (2002). *The brain and the inner world: An introduction to the neuroscience of subjective experience*. Other Press, LLC.

Stewart, J. (2000), Evolution's Arrow Rivett: Chapman Press.

Stewart J. (2008) The Evolutionary Manifesto: Our role in the future evolution of life (ECCO Working Paper 2008-03).

Stewart J. (2010) The Meaning of Life in a Developing Universe, Foundations of Science.

Thagard, P. (1996). *Mind: Introduction to cognitive science*. MIT press Cambridge, MA.

Thaler, R.H. & Sunstein, C.R. (2008), Nudge: Improving decisions about health, wealth, and happiness.

Theraulaz, G., and E. Bonabeau. (1999). A Brief History of Stigmergy. *Artificial Life* 5, no. 2: 97-116.

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences* 4, no. 4: 407–427.

Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Belknap Press.

Turchin, V. (1977). The Phenomenon of Science. A Cybernetic Approach to Human Evolution. New York: Columbia University.

Turchin, V. (1993). The Cybernetic Ontology of Action. Kybernetes, 22, 10-10.

Tye, M. (1986) 'The Subjective Qualities of Experience', *Mind* 95, 1-17.

Valckenaers, P., H. Van Brussel, M. Kollingbaum, and O. Bochmann. (2006). Multi-agent coordination and control using stigmergy applied to manufacturing control. *Multi-Agent Systems and Applications*: 317–334.

Varela, F. J, E. Thompson, and E. Rosch. (1992). *The embodied mind: Cognitive science and human experience*. The MIT Press.

Vidal C. (2007) An Enduring Philosophical Agenda. Worldview Construction as a Philosophical Method (ECCO working paper 2007-03)

Vidal, C. (2008) Wat is een wereldbeeld? (What is a worldview?), in Van Belle, H. & Van der Veken, J., Editors, Nieuwheid denken. De wetenschappen en het creatieve aspect van de werkelijkheid, in press. Acco, Leuven.

Von Bertalanffy, L. (1973). General system theory, G. Braziller.

Von Foerster, H. (1976). Objects: tokens for (eigen-) behaviors. In *ASC Cybernetics Forum*, 8:91–96.

Von Glasersfeld, E. (1996). *Radical constructivism: A way of learning*, Routledge.

Whitehead, A. N. (1978). *Process and Reality: An Essay in Cosmology* Ed. David Ray Griffin and Donald W. Sherburne. Free Press, New York.

Wilber, K. (1997). An integral theory of consciousness. *Journal of Consciousness Studies* 4, no. 1: 71–92.

Wright, R. (2001). Nonzero: The logic of human destiny, Vintage.