

Cognition and (Artificial) Cognitive Systems

- explanatory & exploratory notes -

by Hans-Georg Stork*

Abstract: We briefly sketch some of the many concepts and issues involved in cognition and cognitive systems research and engineering, and describe in very broad terms, some of many potential application areas. We emphasise the need for exploring new architectures for artificial cognitive systems as well as new approaches to modelling and building such systems, also informed by insights into the structure and workings of natural cognitive agents. Future research agendas should be aware of this need while pursuing realistic short- or medium-term goals. These agendas can only be implemented by soliciting contributions from a variety of pertinent disciplines and fields. The few references are meant mainly to illustrate the wide scope of this fascinating emerging area. They are, of course, bounded by the extent that scope is covered by these notes. URLs of preprints are indicated where available.

"It follows that in order to produce consciousness within a mechanical brain, entirely new designs will be necessary. These novel designs will contain as sub-systems the present type of calculator (although in a considerably improved variety) with a very significant additional feature: these sub-systems must perform their own coding. The reason is obvious: as long as man does the coding, the logical principles according to which the calculator is working are located in part outside the machine: they are represented by the actions of the person who does the coding. As long as that is the case, the calculator is not in the possession of vital information (retained by the coder) that is needed to whip the information into proper shape for the transcendental "carrier" operation."

Quoted from: *Can Mechanical Brains have Consciousness?* by Gotthard Günther; Originally published in: *Startling Stories*, Vol.29, no.1, p.110-116, New York 1953

1. Key notions and features

A cognitive agent is a situated entity¹ that is capable of acting ("behaving"), based on evaluating its environment (that may or may not be populated by other cognitive agents) and its own presence therein²; it thus responds to changes in the environment, including those caused by its own actions ("behaviour"), by updating its evaluation of the environment ... etc., ad nauseam: it is caught in an "evaluation - action" loop.

The term "cognition" refers to the intertwined processes underlying the acquisition / building of competences and actionable knowledge by a cognitive agent.

An agent's cognitive capabilities are based on and limited by its intrinsic structures and functions. The precise nature of the structures and functions enabling cognition is not known; exploring these and cognitive capabilities in general, is within the remit of "Cognitive Science".

Cognition emerged in the course of the biological evolution ([8], [28], [29]), mainly in animals, as a specific property of living matter. There appears to be in nature a quasi continuous spectrum of cognitive potential. It culminates in the phenomena of (self-) consciousness, intentionality and the notion of time, supposedly experienced by humans but most likely not only by humans ([12]).

* The views expressed in these notes are those of the author and do not engage his employer.

¹ We call an entity "situated" if it cannot be studied, described or otherwise fully understood, in isolation from some larger context. A pebble in a brook is situated. A situated entity need not be cognitive.

² This is the difference between a fish and a pebble in a brook: the pebble has no way of evaluating (measuring) the flow around it and actively positioning itself accordingly; the fish may even swim against the flow.

No scientifically sound explanation has yet been found of these phenomena and no ways either of measuring the value of an entity's cognitive potential or its degree of consciousness.

There is, however, evidence that a high degree of consciousness (in the intuitive sense of the term) is at least one of the prerequisites for the social and cultural processes that are characteristic of homo sapiens; they hinge on faculties such as language and on symbolic and scientific reasoning which are among the so called "higher cognitive functions". Clearly, these functions and in particular the modes of applying them in specific contexts must have emerged from and be dependent on "lower (or pre-conscious) forms" of cognition.

Yet one may wonder why human consciousness (with all its consequences) seems like a discontinuity in the evolution of cognition. A possible explanation: the apparent gap between the cognitive capabilities of today's homo sapiens and those of its closest living phylogenetic relatives may have been opened up at some stage in the distant past by homo sapiens himself - and his predecessors - who probably eliminated or drove to extinction their less capable competitors.

For an entity to be able to "evaluate its environment" it is necessary that it can "make sense" of it. Hence, there must be a link between the entity and its environment, across which the state of the environment can affect the state of the entity. With natural cognitive agents this link consists of sense organs that are bound to a clearly delimited body. "Making sense" then means to distinguish and identify objects, events and processes in the environment, as well as the relationships between these, and between the entity itself and its environment.

"Perception", the act of "making sense", yields dynamic "representations" associated with the entity, of the system consisting of the entity and its environment. With natural cognitive agents such representations are presumably implicit: they do not take the form of clear cut structures imprinted on some substrate (like text on a piece of paper) but become manifest in changes of the agent's own intrinsic and usually highly complex state space, involving the agent's organism in its entirety. (Is "The 'body' is the representation" a variant of "The medium is the message"?)

Natural cognitive agents do not obtain representations (or "bodies" with all their "ins and outs") as an end in itself. Rather, representations provide the basis for action within their environment: there is no evaluation without representation and both may in fact be considered two sides of the same coin (or "chicken and egg"). A thorough understanding of these processes and the way they interrelate probably requires an equally good understanding of how the biological evolution was primed (or "bootstrapped")³ (see also: [35]). Presumably, an overarching goal arose then: to ensure the stability and durability of certain forms of organic matter, a goal from which natural cognitive agents apparently derive their behaviour (in terms of "making choices for action") in every instance⁴.

It is worth noting that (i) the overarching goal itself in all likelihood has resulted from even "deeper" (physico-chemical) properties of ("non-bio") matter ([13]) and (ii) that the ways and means of generating behaviour have undergone evolutionary change (concurrent with the evolution of "bio-matter"). Specifically human behaviour seems to be, to some extent at least, closely related to the aforementioned phenomena of (self-)consciousness and intentionality.

"Evolution" (or "phylogenesis") denotes but one class of mechanisms by which natural cognitive agents attain their capabilities. These mechanisms run in the very long term, over periods that cover many generations of individuals. Their impact is mainly on the (physical and

³ Indeed, evaluation and action do take place in an organic soup, among molecules that "seek out" structurally fitting companions, to form more or less stable compounds. DNA computing is based on such "behaviour".

⁴ Note that we use the term "goal" in its "neutral" sense; it has neither positive nor negative nor teleological connotations and does not presuppose any deliberate "goal-setting".

behavioural) structures that support expedient representations and operations on them. By contrast, “learning” is closely linked to “ontogenesis” (or a natural cognitive agent's development as an individual, its “growth”); it constitutes, in fact, one of the most salient cognitive capabilities brought forth by evolution, allowing agents to modify their evaluation/representation processes (semi-) permanently as they go along, to create and update their knowledge of “their world”, and to “improve” their behaviour with respect to some “objective function” (of usually very many variables). Homo sapiens appears to be most advanced in this regard as well.

Both, *evolution* and *learning* are ways of *grounding* a natural cognitive agent's behaviour in its environment, of deriving actionable properties, that is, of the objects, events and processes that make up an environment, and of incorporating these properties in the agent's intrinsic state space and state transition mechanisms. Evolution and learning endow natural cognitive agents with a largely unique history.

However, given the “definition” of “cognitive agent” adopted for these notes, grounding is not a process that leaves its imprint on the agent alone: natural cognitive agents adapt to their environment but they may also adapt their environment through their behaviour, in line with whatever goals of their actions. Evolution is in fact co-evolution (of agents and environment) and learning can be co-learning between agents providing there are more than one of them around.

The distinction between the agent and its environment is crucial when talking about cognitive capabilities. It implies that cognitive agents are *open* systems that exchange information and (certainly in the case of natural cognitive agents) energy with entities external to the system, which in turn implies that there must be an interface, a “line of separation” between the agent and its environment. It must be “clear” at any moment in time (or: it must be part of the agent's “knowledge”) what is in, what is out and what belongs to the border.

One may of course ask whether the notion of a “cognitive environment” in isolation has any significance. Clearly, the answer is a resounding “no”: there is nothing a *closed* system could apply its cognitive abilities to and hence the notion is meaningless. This may be a mere corollary to the proposition that any closed system is intrinsically meaningless. The “closedness property” implies that such a system cannot observe or be observed because “observation” is impossible without “interaction” and would therefore violate the “closedness property”. “Meaning”, on the other hand, is only established through observation, interaction, or dialogue. It is the observer that “means”, never the observed. (“Beauty is in the eye of the beholder”.) (It follows for instance that the Universe as a whole is likely to have no meaning whatsoever: the only way to give it meaning is to postulate the existence of a being outside of and interacting with the Universe. But this would contradict the “definition” of “Universe” as encompassing everything. And conjectures to the effect that a closed system may observe itself are at best undecidable.)

2. *Artificial cognitive systems*

Although we have frequently used the qualifier “*natural*” in the preceding paragraphs to illustrate particular points (such as “representation” and “grounding”) we are equally - if not more - interested in *artificial* cognitive agents. And indeed, endowing artefacts (i.e. technical, non-biological things) with cognitive capabilities has been on the agenda of researchers and engineers for many years (centuries, in fact). It has been a dream at first (and occasionally a nightmare), but seems to get closer to realisation as presumably pertinent technologies (electronics and micro-electronics, in particular) are becoming more sophisticated and its products more malleable. Today it is no longer science fiction to speak about and envisage the construction of “artificial cognitive systems”.

There are at least two (equally plausible) reasons for this growing interest:

- (1) the assumption that in order to support and enhance human cognitive activities (including learning) more effectively and to relieve people of the “*cognitive load*” inherent in all sorts of mundane tasks (and daily chores), “machines” should themselves be endowed with cognitive capabilities (and exhibit an appropriate degree of autonomy);
- (2) an expectation that our understanding of natural cognitive systems (and in particular of the cognitive and intellectual faculties homo sapiens - the species - has developed in the course of its biological and socio-cultural evolution⁵) can gain from research into creating artificial cognitive systems and experimenting with them.

2.1 Assessing the computational stance

In the early days of electronic computing and for a considerable period of time, the approach chosen for pursuing the corresponding goals had been guided by the prevailing “Turing paradigm”, leading to a reduction of all things mental to algorithmic symbol manipulation. This stance, frequently referred to as “computationalism”, had been reinforced by the obvious success of digital technologies (i.e. binary switches and devices that set switches, arranged according to the so called “von Neumann architecture”) which, after all, had given rise to many useful tools, extending and amplifying man's own rather limited computing power.

Pure “computationalism” (sometimes - wrongly - equated with “Artificial Intelligence (AI)” research, [36]), however, has not taken us very far yet, towards building artificial systems with cognitive capabilities that would match those of human beings (or other animals, for that matter). It has not contributed much either, towards understanding the principles and architectures underlying these very capabilities.

Two cases in point, chess and natural language, stand out, corroborating this claim:

The game of chess, like any “perfect information” game, admits (at least in principle) a complete “algorithmic solution”([15]). Yet it took surprisingly long (or so it seemed to the proponents of the algorithmic approach to human cognition) to develop the powerful (silicon based) hardware and a computer programme that - albeit not implementing a “complete solution” - defeated the chess world champion in a tournament-style match.

Apparently, human grandmasters have managed to have the upper hand on their digital challengers for almost five decades because their approach to playing chess is largely non-algorithmic or, to put it less boldly, not algorithmic in the sense of the Church-Turing thesis⁶. If it were, humans could not perform certain tasks as well as they (obviously?) do: such as discovering, recognising, evaluating or creating (often fuzzy) complex patterns - not only on chessboards and not only in space, but in time and space-time as well; these tasks can be inherently so compute intensive that they are often beyond the limits of tractability by discrete algorithms (or “symbol manipulation”). And finding the winning move in a game of chess is such a task. (In technical terms (cf. for instance [32]): chess belongs to the complexity class EXPTIME-complete. The game of Go, by the way, also in EXPTIME-complete, has not been mastered yet by computer engineers (of hardware and software), to the extent that their products would defeat a human player of average talent.)

⁵ This is the age-old quest (encapsulated in the pre-Socratic imperative ΓΝΩΘΙ ΣΕ ΑΥΤΟΝ! (“know thyself!")) for coming to terms with our own cognitive capabilities and our behaviour - as individuals, in groups and in societies.

⁶ The Church-Turing thesis claims that for instance Turing machines or - for that matter - von Neumann type digital computers, adequately implement the concept of computability, viz. “recursive function”, viz. obtaining results through manipulation of discrete symbols, based on a finite set of rules that are themselves expressible in terms of discrete symbols. Various attempts (most notably the CHREST project, [45]) to mimic the human approach have so far produced only far from masterly results.

The complexity trap had in fact been well known from very early on: theoretical computer scientists have devoted no small effort to exploring the limitations of the (purely) computational approach to modelling cognition and in particular cognitive tasks considered “high level” (such as reasoning, planning or decision making). They have been extensively studying the combinatorial problems (including problems of formal logic) arising in the context of these models, taking into account deterministic, non-deterministic and stochastic methods ([16]). The results obtained appear to allow but one interpretation: the specific “reasoning” capabilities that distinguish humans from other known forms of life, *when stripped to their symbolically representable (and hence computer implementable) features*, do not suffice to cope with said class of optimisation and decision problems. Worse: implementing these capabilities on digital computers is only of very limited help. Moore's law is no remedy: gains in computing cycles or memory density, governed by its exponential graph, will always be outdone by “combinatorial explosions” which are exponential at best. (See for instance Michael Rabin's paper on “Theoretical impediments to artificial intelligence”, presented at the IFIP 1974 conference, [34].)

At about the same time when computer scientists set themselves the goal of writing a programme that would beat the world chess champion, fellow researchers (including linguists) raised high hopes for automatic translation: within a few years it should be possible to have computers output quality translations of scientific texts for instance from Russian into English. However, while some progress in “machine translation” has been made since, the “quality” issue remains far from being solved ([22]). Apparently, “classical” digital computing cannot cope with many of the features and problems that are peculiar to natural language.

The computational complexity of finding optimal solutions in certain very large search spaces may not be the ultimate culprit in this case. It is, rather, the apparently low degree of possible formalisation that limits the algorithmic tractability of natural language: natural language, reflecting an individual mind's understanding of its environment, is fraught with “analogue components” (of diverse origins, such as history and culture, emotion, situation, etc.) that are simply too vague for being made fully explicit. The difficulty of understanding completely a person's verbal utterances may in fact be of the same tall order as that of the so called “hard problem” of consciousness ([7]): to a large extent, what someone says reflects no less than the private (“subjective”) phenomenal experiences (or “qualia”) of the speaker.

Both examples show that the “power” of human intelligence appears to derive from a characteristic mix of symbolic (“discrete” or “digital”) and “non-symbolic” (“fuzzy” or “analogue”) processes. Humans excel at inventing discrete algorithms of all sorts (with lots of intuition) which is in fact another characteristic feature of human intelligence; yet they have not succeeded in inventing any that would match that intelligence. While the theory and practice of discrete computing are products of human cognitive processes these very processes may themselves not be reducible to or explainable in terms of, the manipulation of discrete symbols. Natural brains can – in principle – do what computers can do but they may not only do that. (One may add: “Why should they? Because they have come up with the idea of a digital computer that outperforms humans, by several orders of magnitude, when executing their algorithms with great precision and at breakneck speed?”) And computers are not brains. Hence, what brains can do, computers (discrete symbol manipulators, that is) need not be able to do. “Matter” (hardware) does matter!⁷ (See also [44])

Our examples represent extremes within the range of “high-level” cognitive tasks modern man can or has to face up to. Proponents of the purely computational approach to cognition have succeeded on the chess front but they have done so using methods that are quite different indeed from the ones human chess masters appear to apply to finding the “winning move”.

⁷ Contrary to the functionalist's claim that cognitive capabilities are independent of their physical substrate.

Chess happened to be amenable to the formalisms they have at their disposal, and not surprisingly so, given that it has been *invented* - by “informal brains” (!) - as a formal game in the first place. On the other hand, they have failed on natural language because it may be too difficult a domain to model, beyond some comparatively simple formal structures and statistical properties, i.e. some more or less *abstract* features. (One may add: “This is no surprise either, given that natural language has somehow *emerged* within interacting human brains; it has not been invented by them!”)

Consequently, critics of computationalism have not only picked up on the issue of computational complexity but also on the difficulty of capturing through explicit symbolic representations, the subtleties of “living structures”, such as natural language. The main concern is about “semantics”. While the rules of games such as chess mean nothing beyond “what they are”, words and expressions in natural languages have meanings other than these words and expressions themselves. To put it differently: for formal games such as chess, syntax and semantics are identical, whereas the syntax of a natural language, regardless of its level of detail, does not appear to convey the full semantics of that language. The same holds for any formalism designed with a view to creating abstract (“detached”), explicit representations of real-world phenomena.

It boils down to two fundamental and interrelated problems that preclude declaring programmes (and their supporting systems, both software and hardware) “cognitive” if they are based solely on externally imposed representations:

(1) No such representation is entirely adequate. This is also known as the “frame problem” (or “*ceteris paribus*” problem, [33]): to focus on what needs to be known (and represented) about a given environment and to ignore what can be safely ignored. For most real-world environments it is hardly possible to make this distinction at design time. A robot for instance that is given the floor plan of a building, formal descriptions of the objects it may encounter, and rules to govern its movements and other actions, would be at a complete loss if slight changes, not covered by the instructions given, were made in its environment through its own actions. Likewise, a programme that is supposed to react sensibly to arbitrary statements made in some natural language (for instance by translating them into another language) would require access to an appropriately comprehensive (yet manageable!) model (in terms of some explicit representation) of the world(s) that this language is (or these languages are) about.

(2) Such representations would only reflect the understanding their human designer has of the things they are supposed to represent. They would be “grounded” in what he or she can consciously conceive of; they would not have any semantics (meaning) “of their own accord”. (In other words: their semantic ground would be of the same quality as that of any odd manmade text which may or may not represent the knowledge - or ignorance - of its author.) This is also known as the “(symbol or representation) grounding problem” ([19]) which has been brought to the attention of relevant research communities through John Searle’s famous “Chinese Room” gedankenexperiment ([38]), sparking a debate that has been going on for more than twenty years now (cf. for instance [18]). While this debate has touched on elusive philosophical issues such as the notion and nature of “understanding” or the “mind-body” problem, perhaps its greatest impact has been on rethinking the relationship between syntax and semantics: The person in the “Chinese Room” is given strings of Chinese characters (that are perfectly meaningful to the people outside) and produces intelligible and sensible replies – also in Chinese – not based on her “knowledge of the language” (in fact she has none whatsoever) but (and that is the point) exclusively on a set of syntactic rules she has at her disposal. That person does not learn anything about the meaning of the Chinese characters she is manipulating (for instance of the symbols that to people outside represent the concept of a Beijing duck), beyond the way they are to be manipulated, because she has never *experienced a connection*

between these characters and something in the outside world⁸. Searle concludes that when it comes to “real-world” matters abstract (or “detached”) syntax alone need not yield any semantics at all (contradicting John Haugeland’s claim: “If you take care of the syntax the semantics takes care of itself.” [21]).

To put it in a nutshell, these problems are about the breadth (of coverage of relevant situations) and depth (of the semantic ground) of actionable representations. They are fundamental indeed as every living creature and every species have to solve them during their respective lifetimes. Natural cognitive agents (as species and/or individuals) do develop adequate representations with intrinsic meanings. (Here, “adequacy” means “sufficiency for survival”, at least in the short and medium term; “perfection”, of course, is not a requirement!) As pointed out above they do this by employing mechanisms (such as evolution and learning) that allegedly entail the adequacy and semantic depth of the representations they work with: proper grounding may in fact also yield solutions to the “frame problem” ([20]). (Note that Searle, oddly enough, in his gedankenexperiment, did not seem to question the possibility of a full syntactic representation of all the knowledge needed to produce answers in Chinese that are meaningful to the people outside the “Chinese Room”!)

The upshot of the discussion so far appears to be:

- (A) “Classical” Artificial Intelligence programmes give only the appearance of replicating high level human cognitive capabilities if the semantics of the objects they deal with is entirely externally defined; they are no more “cognitive” than programmes that help engineers solve the odd differential equation.
- (B) An artefact cannot be deemed “cognitive” if it does not derive the meaning of the objects it is supposed to deal with (or, equivalently, its knowledge, beliefs or assumptions about these objects and what to do with them), from largely autonomous (in the sense of “un-supervised”) interactions with these objects; it should not be grounded externally (by design); it should ground itself.

Two key questions follow: if artificial cognitive systems are fundamentally different from (classical) computing systems then

- (i) what should they do that (classical) computing systems presumably cannot do, and
- (ii) how should they be constructed?

2.2 What should artificial cognitive systems do?

An answer to (i) should certainly add to and further elaborate on the reasons given above for our growing interest in artificial cognitive systems, viz. their potential for enhancing people’s own cognitive capabilities and lessening human cognitive load. A first clue may be provided by looking at current computing machinery, the kind of equipment many people are in one way or another involved with at their workplace or as ordinary citizens. Whether *stand-alone*, *networked* or *embedded*, these machines and systems suffer (to a greater or lesser extent) from a number of rather nasty shortcomings, e.g.:

⁸ For this very reason popular attempts at debunking the Chinese Room argument by conceding “understanding” to the whole room (rather than to the person inside), are flawed: the entire room would not have any idea either, of what the characters for “Beijing duck” mean, if it had not had the experience of seeing, smelling, touching or tasting such an animal - and made some internal representation thereof. The only “experience” the room gains is that of the contexts (in terms of other characters) in which characters appear. Of course, the totality of these contexts does constitute “a world” of which “the room” may even build internal representations (an “ontology of contexts”) beyond the syntax rules it has been endowed with. But that world is not the “real” world which our senses connect us to. It is as formal a world as the one spanned by the rules of chess. The room would be engaged in a “sprachspiel” (à la Wittgenstein, [49]) at best, without any reference to an outside reality.

- they crash or make their host-systems crash or dysfunctional; they give cryptic error messages or “withdraw” for no apparent reason;
- they do not easily adapt of their own accord to those who use them, for instance by taking advantage of individual usage histories;
- many of them require deep expertise and considerable effort to configure, update, debug, repair and protect;

which makes it difficult and expensive to take full advantage of their otherwise astounding capabilities.

Given our growing dependence on information processing devices in almost every walk of life, and the cost this entails, these shortcomings alone seem to justify considerable effort to enable such systems ...

- to adapt themselves to their users' demands and idiosyncracies (including emotional ones), and – within reasonable limits - to changing overall requirements;
- to become proactive “partners” in interactions with people;
- to deal robustly with failure and – more generally - act sensibly upon encountering unforeseen situations;

by:

- learning about their environments and their own situation therein, and by making their reactions and actions contingent on what they learn;

that is, in short, by

- exhibiting at least some of the qualities usually ascribed to natural cognitive systems.

Moreover, apart from supporting people in all sorts of mundane tasks (including their daily chores, e.g. by relieving them of the “cognitive load” inherent in such tasks and owing to the increasing complexity of man’s own inventions), artificial systems with cognitive capabilities are desirable or needed to ...

- help unlock environments that are not normally or not easily accessible to humans, e.g.:
 - hazardous or hostile environments, but also
 - environments that would require, to be fully understood and exploited, sensory capabilities humans do not possess (but which can be provided by artificial sensors and measuring devices); or
 - artificial, yet evolving, environments, such as complex (man-made) systems and networks (this includes not only “computer networks” and their contents, but also “traffic/transport networks” of all kinds!),

and thus enhance and expand people’s own cognitive capabilities (beyond mere amplification or support of symbolic reasoning and calculation);

- accelerate the efficiency of human learning, e.g. through intelligent teaching devices that could enable a qualitative leap in the ability of humans to acquire skills and understanding.

Concrete applications do of course depend on the characteristics of the environment a system is supposed to operate in. In principle, there is no limit. Generally speaking, artificial cognitive systems might play a major role in all environments where ...

- robustness of perception (e.g. through recognition, analysis and *understanding* of patterns in space, time and space-time) and (re-)action, and
- the ability to acquire and organise relevant knowledge and to “cope with change”

are critical success factors. Depending on the frequency of change this may entail more or less strong real-time requirements on system reaction.

Systems could co-operate with people in creating and updating ontologies (in whatever form deemed suitable) pertaining to their respective environments or social organisations, thus turning “socio-technical systems” into “hybrid cognitive systems”.

It follows that the remit of artificial cognitive systems and/or their underlying technologies may include (to name but a few):

- remote and on-site environmental sensing and monitoring,
- management and control of energy or communication (e.g. road or computer) networks,
- medical diagnostics,
- vehicle control and traffic safety,
- industrial manufacturing,
- metadata production in large (distributed) repositories of digital content;

apart from spawning new developments in ergonomics and improving ways of using systems that mediate transactions based on all sorts of digital content (e.g. ticket machines).

Clearly, to achieve the above objectives it may not be necessary or even desirable to emulate or instantiate the full range of human cognitive capabilities (this may, at any rate, not be possible and incur the danger of being ridiculed - as the presumably functionalist⁹ approach of “strong AI” has been ridiculed at times by some people). On the contrary: artificial systems should have cognitive capabilities only to the extent of being more responsive to human demands, and they should outperform people in areas where human capabilities are limited.¹⁰ However, artificial systems - as “cognitive” as they may be - must not be allowed to “take over”. (For a discussion of “human likeness” see for instance [41])

2.3 Creating artificial cognitive systems – challenges and issues

Turning to the question of how artificial cognitive systems can or should be constructed we have to contend that it seems entirely unfeasible to start from scratch as nature did. But we can learn from nature as generations of engineers did who designed their products based on insights gained in physics and other “classical” sciences.

The “methods” used by nature do after all, yield the structures and representations that support cognitive processes, including those that underlie the “higher cognitive functions” enjoyed by humans.

The *main challenge* appears to be in finding ways of *adequately grounding* a system’s structures and representations in whatever environment that system is supposed to operate in (cf. our conclusion (B) above and [51]). This entails establishing close links between learning, evaluation (through perception and reflection) and action (including action upon the system itself). A key problem (if not the key problem) would be to determine “where design ends and (semi-) autonomous evolution, self-organisation and learning begin”.

Taking this decision and following up on it is likely to call for approaches that go way beyond and may in fact be fundamentally different from, the methods and techniques (e.g. for “high-level” reasoning, planning and problem solving in general) that are usually being associated with “classical” symbolic Artificial Intelligence (cf. our chess and natural language examples).

⁹ cf. footnote 7

¹⁰ Indeed, this is trivial: given the differences between natural and artificial hardware (“matter” matters!!) machines do certain things better than humans and vice versa; just like a well-trained dog is better than its trainer at sniffing drugs in a bag.

In all fairness we must admit though, that parallel to largely “meta-mathematics inspired” AI research (that appeared to dominate the agenda for a considerable period of time in the previous century) there have always been “engineering oriented” and “biology motivated” strands of activities. While definitely picking up on issues pertaining to cognitive systems these strands were often not explicitly associated with “mainstream” AI. Both had in fact once (in the 40s and 50s of the last century) been intertwined in a discipline called “Cybernetics”, a term coined by Norbert Wiener to denote the science of “communication and control in animal and machine” ([48]). It had certainly been an early contender in the race for artificial intelligence, yet fell behind, presumably because its hardware requirements could not be met when electronics was still in its infancy.

But the flame has been kept burning. And in light of the fact that natural cognitive agents (as individuals or species) are (up until now) practically the only entities that are capable of learning through acting on or interacting with their environment, it now seems obvious that the engineering of artificial cognitive systems should be informed by studying natural processes related to cognition and control, and their physical substrates.

Especially during the last two decades of the previous century, this insight has made, in one form or another, a significant impact on cognitive science and systems research; it has led to the revival of various lines of work that had been relatively dormant for some time, to wit: on machine learning and artificial neural networks (or, more generally, on parallel distributed processing (PDP) architectures ([1], [30])) and their application to pattern recognition and classification and, more specifically, to vision and natural language understanding, among others.

Machine Learning ([10]) and Artificial Life ([5]) research had in fact long since picked up on “nature's ways”: genetic algorithms for instance, and other variants of evolutionary procedures have been studied in the early seventies of the previous century, albeit not in the mainstream of computer science (and often not even with a view to applications of computing). Strategies gleaned from the biological evolution are now also being utilized for the development of both, software and hardware, yet not on a large scale ([25]). And evolving networks (of digital content and services, such as the World Wide Web) have become popular domains for applying learning algorithms to improving the usability and searchability of these networks (e.g. through “ontology learning” ([27]) and “emergent semantics” ([17])).

Indeed, “learning” appears to be the “Eternal Golden Braid” that ties together natural and artificial cognitive systems: for the former it materialises in the genetic (phylogenesis) and the neural (ontogenesis) substrates of an individual's cognitive performance; for the latter, it is the crucial ingredient, and suitable forms are being researched and developed, based also on nature's own inventions. It may not be too far fetched to predict that “training”, “teaching” and “learning” – rather than explicit and external “programming” – will eventually empower systems whose functioning depends on “keeping in touch” with some environment that is part of or connected to the real world, and on expedient information processing.¹¹

One of the key issues (and a corollary to the “grounding challenge”) is to find suitable architectures for evolution and learning – for building and updating in largely autonomous fashion the structures and representations that support the actions and reactions of a cognitive system, as required in a given environment. Both, hardware and software are concerned; ideally, they would have to allow for permanent modification (“self-organisation” and “self-construction”) in response to changes within the environment.

¹¹ Alan Turing himself, one of the founding fathers of Artificial Intelligence and Cognitive Systems research, in his seminal paper on *Computing Machinery and Intelligence* ([42]), formulated this as part of a future research programme aimed at achieving “human-like” machine intelligence.

This will entail rethinking the very concepts of hardware and software. Already today the distinction between hardware and software gets blurred for instance in field-programmable gate arrays (FPGA, [31]) or cellular neural/nonlinear networks (CNN, [11]). Incidentally, animal brains and organisms do not make that distinction at all. Presumably, their “software” (and data!) consists in *the way their hardware components (neurons and other cells) are arranged and linked*; it “executes itself” through rearrangement of components and modification of link attributes; memory and processors are not separated as in the classical *von Neumann* computer architecture.¹²

It follows that

- an increase in sheer processing power and memory capacity alone, of our current computing devices, even if pushed to their physical limits, will probably not yield the desired results; and that
- research will have to take into account or even initiate (which would be yet another major challenge) new hardware developments.

Recent advances in the neurosciences that lead to a better understanding of “brain dynamics”, may have a significant impact. Results for instance, on how brains solve the “binding problem” (e.g. of sensory features, [14], [40]), or on how natural (e.g. associative) memory mechanisms work ([23]), may be of great relevance.

Another – somewhat closer - source of inspiration as far as “soft” configuration of hardware is concerned may also be found in technologies that belong to the “hard core” of current computer engineering: fault-tolerant computing for instance, large scale packet-switched computer networks and “Grids”. Packet-switched “networks-in-the-large” in particular, are paragons of fault-tolerant and adaptive systems: they degrade gracefully upon failure of nodes and can allocate resources dynamically, depending on load and other demands. They also upgrade (“grow”) gracefully upon adding nodes ([24]). Thus, while still being firmly rooted in human social structures and not growing or evolving “of their own accord” (but also not grounded in premeditated design only) they do exhibit important characteristics of living organisms and brains (!): “plasticity” for instance and the ability of limited self-organisation. Protocols and mechanisms similar or analogous to packet-switching “in the large” exist that support fault-tolerance and sustain self-organisation in small(er) networks (or very small ones, be they called “neural” or whatever else) of sensors, transmitters and actuators ([2]).

The search for cognitive architectures that are suitable for and feasible in a given (type of) environment will also give rise to rethinking the concepts of embodiment and physical instantiation of cognition. In fact, there is no information processing (and in particular, no computation) without “a physical body”; processing information and computing do not happen in the void, they are always bound to physical substrates which are key to the understanding of both qualitative and quantitative aspects of information processing.¹³ Hence, the type of “physical instantiation” of a specific cognitive capability becomes irrelevant only if one adopts the “functionalist” stance. Simulation may help greatly to gain valuable insights about all sorts of phenomena but it is never the “real thing”.¹⁴

Answers to the question “what kind of body does a cognitive system need?” must take account of the particular physical features of the system’s environment and of the role the system should assume in it (cf. the characterisation, given further above, of candidate environments for artificial cognitive systems). As pointed out before, the ability to distinguish be-

¹² This, of course, is what artificial neural networks (ANNs, [1]) are supposed to model.

¹³ To follow up on footnote 10: mechanical calculators, DNA or (as yet hypothetical) quantum computers are quite different from “silicon-based” computers as regards these aspects.

¹⁴ Cf. footnote 7. A “radical functionalist” would probably proffer the possibility of instantiating (or re-instantiating) an entire brain without loss of functionality by simulating its processes on some machine.

tween what happens inside a cognitive system and what is going on outside is all-important. To make that distinction it may not at all be necessary to endow the system with a physically discernible, bounded body (to put it “in a box” or on legs). It needs to be clear though, which processes are “in” (i.e. belong to the system), which are “out” (i.e. do not belong to the system but to its environment), and how (via what channels) and what the processes belonging to these two classes respectively, communicate (this property is occasionally referred to as “structural coupling”, cf. [50]). The possibility of having cognitive systems work in dynamic data (or “digital content”) spaces must therefore not be excluded¹⁵.

Abstract models of cognitive information processing may have to be devised beyond the classical models, Turing machines and equivalent formalisms that is, of discrete symbol manipulation. Indeed, while the “Turing paradigm” yields appropriate abstractions of the symbolic faculties we likely owe to the phenomenon of consciousness¹⁶, it apparently fails to explain for instance the ability of our (and other animals’) un- or subconscious systems to produce optimal or near-optimal responses fast, to changes in or stimuli from our environment.

It is therefore no surprise that the practical relevance of the Church-Turing thesis (cf. footnote 6) has been put in serious doubt lately, and not only for cognitive information processing. It is challenged from at least four (partly interrelated) perspectives:

- software engineering: the Turing paradigm may be inadequate for studying and reasoning about interactive and networked (software-)objects (Wegner, [46]); there is a need to extend the thesis to cover “*non-uniformity of programs, interaction of machines, and infinity of operation*” (van Leeuwen & Wiedermann, [43], [47]);
- theoretical computer science (complexity theory): computing over the reals is likely to be more powerful than computing over discrete domains (Blum, Shub & Smale, [3]);
- artificial neural networks (ANN): the inclusion of analogue components (analogue recurrent neural networks, ARNN) leads to provable “super-Turing” behaviour (Siegelmann, [39]);
- quantum computing: the physical character of every computational process becomes paramount and the Turing model becomes inadequate (while not necessarily invalidating the Church-Turing thesis) (Deutsch, [9]; Calude shows that, in theory, quantum computers can go beyond the “Turing-limit”, i.e. solve problems not solvable by any Turing machine, [6]).

New non-standard and more “realistic” conceptualisations of information processing¹⁷ are likely to become part of the theoretical underpinning of future cognitive systems research and cognitive engineering. They would have to reflect the different modes of interaction between a cognitive system and its environment (which may include other cognitive systems and people), the unpredictability of external events as well as the largely analogue nature of most real-world processes (including man-made real-world processes)¹⁸, and their concurrency. Subsumption architectures (as introduced by Rodney Brooks in robotics, [4]) and the much discussed phenomena of “emergence” (of properties of systems assembled - through self-organisation - from many small components) and “self-creation” (autopoiesis, [28]) may also fall within the remit of such models.

¹⁵ It certainly has not been excluded in the past as amply documented by a substantial body of research under the heading “intelligent (information) agents”.

¹⁶ Presumably, this is exactly what Turing had in mind when he proposed his formalism.

¹⁷ “Journeys in Non-Classical Computation” is one of seven current UK “Grand Challenges” in Computing Research (<http://www.cs.york.ac.uk/nature/gc7/index.htm>)

¹⁸ Whether or not the real world is best modelled as discrete at its smallest scale is probably still an open question.

An agenda for cognitive systems research and cognitive engineering would certainly have to be aware of the “soft hardware (or hard software)” issue (with all its implications: the ability to evolve, grow, adapt, learn, self-check, self-repair, etc., perhaps even self-replicate). However, viable solutions pertaining to this issue, while definitely on the horizon, are not necessarily imminent. They would therefore constitute rather long-term goals. This begs the question: what could be intermediate milestones and what methods, tools and practical achievements could be linked to them?

2.4 Creating artificial cognitive systems – computationalism regained?

Some answers at least may be found by going back to the “Chinese Room”. Upon closer inspection Searle’s argument may not be as damning as it appears, to the possibility of endowing artefacts with cognitive capabilities by taking a computational approach to processing information, no matter where that information arises, in the “real world” or elsewhere. The gist of the argument¹⁹, as pointed out above, is the fact that the “Chinese Room” (or whatever or whoever is in it) has no way to analyse (or, more generally: to interact with) its environment and derive its own (intrinsic) rules and representations from that analysis (or interaction). All it can do is some sort of pattern matching to find the right rule to apply to the symbols it receives through its input channel. Searle’s critique, while undeniably making an important point in the syntax-semantics debate, has a very strong “batch-processing flavour” about it - as if real-time systems had never been invented and widely used.

Real-time control systems (descendants by the way, of early Cybernetics) owe their very existence to the fact that computing devices can be connected to real-world processes, that they can evaluate what is happening “out there”, and that they can trigger actions to follow evaluation. Examples abound, large (e.g. in factories or power plants) and small (e.g. in aeroplanes, cars or human hearts). Quite popular these days and potentially life saving are microprocessors that are hidden under the hood of a car, connected to sensors that monitor the momentum of the wheels, and to effectors that adjust brake pressure almost instantly should an imbalance be detected, thus preventing the car from skidding. Many real-time control systems (embedded in all kinds of appliances) have limited learning capabilities that allow them to change state transition probabilities depending on a particular history (a time-series of measured values, for instance).

Examples such as these may indeed qualify as primitive artificial cognitive systems, not unlike primitive life forms qualify as natural cognitive systems. They stand for what might be called “environment-enriched” computationalism. Unlike natural language understanding programmes, based on scripts (once popular symbolic “knowledge representation” devices proposed by Roger Schank ([37]), and the original targets of Searle’s argument), they are not detached from their environment. Hence, while still being “grounded in design” and abiding by rules of discrete symbol manipulation, they do show us a way out of the “Chinese Room”.

The crucial components of real-time control systems are - almost literally - eyes, ears and hands: sensors and actuators, that is. The degree of interaction with the environment made possible by these components, and the effect of interaction on the intrinsic state space and state transition function of such systems, determine their “cognitive power”. Existing or about to emerge hardware and software technologies already offer many possibilities to implement sophisticated schemes for interaction-based representation-building and learning; developing these schemes further and applying them may eventually turn real-time control systems into fully fledged embedded cognitive systems. For instance into one - to pick up on the above example - that makes a vehicle even safer by integrating not only stabilisers and anti-blocking

¹⁹ ... or perhaps its basic flaw!

mechanisms but also surround vision and an episodic road-memory that would allow the car to react to unexpected objects that are in its way or to anticipate possibly dangerous bends. Clearly, the goals guiding the evolution and learning of such systems should be to avoid fatal collisions and prevent injury and damage²⁰, as foretold (some forty years ago!) for instance in [26].

A system would, no doubt, be considered endowed with “high-level” cognitive capabilities, if it could interpret scenes where objects (including the system’s “body” itself) move in three dimensions, if it could look ahead and foresee potentially dangerous situations, and if it could act on percepts and plans – either directly or by alerting a human operator or monitor and offering suitable choices for action or decisions to be taken.

Of course, not all environments and not all cognitive tasks involve or need real-time feedback and control. Also, the type of possible feedback (from humans, physical processes or other cognitive systems) and the tightness of control may vary greatly depending on the given environment and task (cf. section 2.2). Yet different application scenarios exhibit many commonalities.

Cognitive systems research would have to *focus on these commonalities* and in the short and medium term propose generic or specific solutions through “state-of-the-art” implementations of “higher” cognitive functions (perception, knowledge construction, reasoning and communicating in human terms), based on “*low-level*” *real-time interaction* or – in the absence of strict real-time constraints (as in most “digital spaces”) – analysis of “*low-level*” *features / signatures* of the system’s environment.

First and foremost research would have to address the issue of learning. And it would have to leave the door open for gradually improving cognitive capabilities (in particular learning) through the adoption of more malleable hardware-software “matter” as it becomes sufficiently mature. The “*sprachspiel*” (“language game”) alluded to in footnote 8 must become a “*reality-enhanced sprachspiel*” (or “serious business”). To some extent simulations will be necessary, possible and useful.

Research also has to address the creation of a firm methodology for the specification of artificial cognitive systems (“what should be the desired capabilities and functions, and how should they be described?”), their design (“what are the most suitable design principles, supporting architectures, where does ‘design end and autonomy begin’, etc.?”) and testing (“to what extent is cognition measurable?”). All this would be based in large measure on “non-standard” information processing models (as discussed above).

Clearly, this research (and subsequent development) needs to be interdisciplinary, drawing heavily on cognitive science (itself a “transdiscipline” with contributions most notably from psychologists, neuroscientists, AI experts, developmental biologists, ethologists, anthropologists, and philosophers of mind), general computer science and engineering, industrial engineering (e.g. flexible manufacturing systems and mechatronics) and, last but not least, mathematical systems theory. Existing lines of research will have to be extended and connected before the emerging mesh can be fleshed out.

²⁰ Research with a view to *providing the methodological and architectural basis* for building “cognitive vehicles” would well be worth an entire funding programme if such vehicles, widely used, could significantly reduce road casualties. Obviously, this - as many other kinds of - application of “cognitive control” is of “dual use”. The DARPA “Grand Challenge” in 2004 for instance, required “autonomous robotic ground vehicles to successfully navigate a course from Barstow, CA to Primm, NV”, in a “quest to develop a new generation of autonomous robotic ground vehicles that some day soon will save the lives of men and women in our armed forces by performing hazardous tasks on the battlefield”.

References

Most of the below references are to survey papers, “popular science” publications or textbooks. They should be easily accessible, both literally and figuratively speaking, with the exception perhaps of some that go into greater technical detail and/or require – to be fully appreciated – a more mathematically tinted background. It goes without saying that this selection represents but a micro-micro-sample of past and current research that bears on our understanding of cognition and cognitive systems. A slightly more comprehensive collection of resources (mainly thematic portals) will be added to a forthcoming version, in HTML, of this article.

- [1] Abdi H. (1994) A neural network primer. *Journal of Biological Systems*, 2, 247-281 (<http://www.utdallas.edu/~herve/abdi.primer.pdf>)
- [2] Akyildiz I.F., W. Su, Y. Sankarasubramaniam, E. Cayirci (2002) Wireless sensor networks: a survey. *Computer Networks* 38 (2002) 393–422 (<http://www.ece.gatech.edu/research/labs/bwn/sensornets.pdf>)
- [3] Blum L. (2004) Computing over the reals: where Turing meets Newton. *Notices of the AMS*, October 2004 (<http://www-2.cs.cmu.edu/~lblum/PAPERS/TuringMeetsNewton.pdf>)
- [4] Brooks R.A., C. Breazeal, M. Marjanovic, B. Scassellati, M. Williamson (1999) The Cog project: building a humanoid robot. in: *Computation for Metaphors, Analogy, and Agents*. C. Nehaniv (ed), *Lecture Notes in Artificial Intelligence* 1562. New York, Springer, 52–87, 1999 (<http://www.ai.mit.edu/people/brooks/papers/CMAA-group.pdf>)
- [5] Brooks R.A. (2001) The relationship between matter and life. *Nature*, Vol. 409, January 18, 2001, pp. 409–411 (<http://www.ai.mit.edu/people/brooks/papers/nature.pdf>)
- [6] Calude C. S., B. Pavlov (2002) Coins, quantum measurements, and Turing's barrier, *Quantum Information Processing* 1, 1-2 (2002), 107–127 (<http://www.cs.auckland.ac.nz/%7Ecristian/coinsQIP.pdf>)
- [7] Chalmers D. J. (1995) Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2(3):200-19 (<http://jamaica.u.arizona.edu/~chalmers/papers/facing.html>)
- [8] Darwin C. (1859) *The Origin of Species*. Gramercy Books, 1998 (<http://home.tiscalinet.ch/biografien/sources/origin.htm> or <http://onlinebooks.library.upenn.edu/webbin/gutbook/lookup?num=2009>)
- [9] Deutsch D., A. Ekert, R. Lupacchini (2000) Machines, logic and quantum physics. *Bulletin of Symbolic Logic* 3 3, September 2000 (<http://xxx.lanl.gov/abs/math.HO/9911150>)
- [10] Dietterich T. G. (1997) Machine-learning research - four current directions. *AI Magazine*, Winter 1997 (<http://www.aaai.org/Resources/Papers/AIMag18-04-010.pdf>)
- [11] Dogaru R. (2003) *Universality and Emergent Computation in Cellular Neural Networks*. World Scientific Publishing Co., 2003
- [12] Edelman Gerald M., Giulio Tononi (2001) *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books, 2001
- [13] Eigen M., R. Winkler-Oswatitsch, P. Woolley (1996) *Steps Towards Life: A Perspective on Evolution*. Oxford University Press, 1996
- [14] Engel A.K. and W. Singer (2001) Temporal binding and the neural correlates of sensory awareness. *Trends in Cogn. Sci.* 5(1): 16-25 (http://www.cognitive-science.net/downloads/paper_engel01a.pdf)
- [15] Fraenkel A. S. (2000) Combinatorial games: selected bibliography with a succinct gourmet introduction. *The Electronic Journal of Combinatorics*, Dynamic Survey #2 (<http://www.combinatorics.org/Surveys/ds2.pdf>)
- [16] Garey M. R., D. S. Johnson (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman & Company, 1979
- [17] Grosky W.I., D.V. Sreenath, F. Fotouhi (2002) Emergent semantics and the multimedia Semantic Web. *ACM SIGMOD Record*, Dec 2002 (<http://lsdis.cs.uga.edu/SemNSF/SIGMOD-Record-Dec02/Gorsky.pdf>)

- [18] Harnad S. (1989) Minds, machines and Searle. *Journal of Theoretical and Experimental Artificial Intelligence* 1: 5-25
(<http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad89.searle.html>)
- [19] Harnad S. (1990) The symbol grounding problem. *Physica D* 42: 335-346
(<http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad90.sgproblem.html>)
- [20] Harnad S. (1993) Problems, problems: the frame problem as a symptom of the symbol grounding problem, *Psychology*: 4,#34
(<http://psycprints.ecs.soton.ac.uk/perl/local/psyc/makedoc?id=328&type=xml>)
- [21] Haugeland J. (1985) *Artificial Intelligence: The Very Idea*. The MIT Press, Cambridge, MA, 1985
- [22] Hutchins W.J. (2001) Machine translation over fifty years. *Histoire, Epistemologie, Langage*, Tome XXII, fasc. 1 (2001), p.7-31 (<http://citeseer.ist.psu.edu/578882.html>)
- [23] Kandel E. R. (2000) The molecular biology of memory storage: a dialog between genes and synapses. *Nobel Lectures in Physiology or Medicine 1996 – 2000* (Hans Jornvall, ed.), World Scientific Publishing Co., 2003 (<http://www.nobel.se/medicine/laureates/2000/kandel-lecture.pdf>)
- [24] Kleinrock L. (1964) *Communication Nets; Stochastic Message Flow and Delay*, McGraw-Hill Book Company, New York, 1964. Reprinted by Dover Publications, 1972.
- [25] Koza J. R., F. H. Bennett III, D. Andre, and M. A. Keane (1999) *Genetic Programming III: Automatic Programming and Automatic Circuit Synthesis*. Morgan Kaufmann, 1999
- [26] Lem S. (1961) *Return from the Stars*. Harcourt 1980
(<http://www.cyberiad.info/english/dziela/powrot/powrotpl.htm>)
- [27] Maedche A., S. Staab (2001) Ontology learning for the Semantic Web. *IEEE Intelligent Systems*, vol. 16, no. 2, pp. 72–79. (<http://www-etsi2.ugr.es/depar/ccia/mabd/material/adicional/semantic/semantic8.pdf>)
- [28] Maturana H., F. Varela (1987) *The Tree of Knowledge: A new look at the biological roots of human understanding*. Shambhala/New Science Library, Boston, 1987
- [29] Mayr E. (2001) *What Evolution Is*. Basic Books, 2001
- [30] McClelland J. L. and Rumelhart D. E. (1988). *Explorations in Parallel Distributed Processing*. MIT Press 1988
- [31] Meyer-Baese U. (2004) *Digital Signal Processing With Field Programmable Gate Arrays (Signals and Communication Technology)*. Springer Verlag, 2004
- [32] Papadimitriou C. H. (1993) *Computational Complexity*. Addison-Wesley Pub Co, 1993
- [33] Pylyshyn Z.W., ed. (1987) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Ablex Publishing Corporation, Norwood, NJ, 1987
- [34] Rabin M. O. (1974) Theoretical impediments to Artificial Intelligence. *IFIP Congress 1974*: 615-619
- [35] Ruiz-Mirazo K., A. Moreno (2004) Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life* 10: 235–259 (https://mitpress.mit.edu/journals/pdf/artl_10_3_235_0.pdf)
- [36] Russell Stuart J., P. Norvig (2003) *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2003 (<http://aima.cs.berkeley.edu/>)
- [37] Schank R., R. Abelson (1977) *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Lea, 1977
- [38] Searle J. R. (1980) Minds, brains and programs. *Behavioral and Brain Sciences* 3: 417-424
(<http://www.bbsonline.org/documents/a/00/00/04/84/bbs00000484-00/bbs.searle2.html>)
- [39] Siegelmann H. T. (1999) *Neural Networks and Analog Computation: Beyond the Turing Limit*. Birkhäuser Boston, 1999
- [40] Singer W. (1999) Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24: 49-65 (<http://www.mpih-frankfurt.mpg.de/global/Np/Pubs/review.pdf>)

- [41] Sloman A. Architectural requirements for human-like agents both natural and artificial (What sorts of machines can love?). in: Human Cognition and Social Agent Technology (Dautenhahn, ed.), John Benjamins Publishing Co, 2000 (<http://citeseer.ist.psu.edu/245650.html>)
- [42] Turing A. M. (1950) Computing machinery and intelligence. *Mind* 49: 433-460 (<http://cogprints.ecs.soton.ac.uk/archive/00000499/00/turing.html>)
- [43] van Leeuwen J., J. Wiedermann (2001) The Turing machine paradigm in contemporary computing, in: B. Enquist and W. Schmidt (Eds), *Mathematics Unlimited - 2001 and Beyond*, Springer-Verlag, 2001, pp. 1139-1155 (<http://archive.cs.uu.nl/pub/RUU/CS/techreps/CS-2000/2000-33.pdf>)
- [44] Varela F. J. (1998) Le cerveau n'est pas un ordinateur: on ne peut comprendre la cognition si l'on s'abstrait des incarnations. Interview by Hervé Kempf. *La Recherche*, No. 308, Avril 1998, pp. 109-112 (http://web.ccr.jussieu.fr/varela/press_releases/LaRecherche308.html)
- [45] Waters A., Gobet, F., & Leyden, G. (2002) Visuo-spatial abilities in chess players. *British Journal of Psychology*, 30, 303-311 (http://www.psychology.nottingham.ac.uk/staff/Fernand.Gobet/preprints/Visuo-spatial_abilities.doc)
- [46] Wegner P., D. Goldin (2003) Computation beyond Turing Machines. *Communications of the ACM*, April 2003 (<http://www.cse.uconn.edu/~dqg/papers/cacm02.rtf>)
- [47] Wiedermann J. (2004) Characterizing the super-Turing computer power and efficiency of classical fuzzy Turing Machines. *Theoretical Computer Science*, Vol. 317, 2004, pp. 61-69
- [48] Wiener N. (1948) *Cybernetics or the Science of Communication and Control in the Animal and Machine*. Hermann Editions in Paris; Cambridge: MIT Press, Wiley & Sons in NY, 1948
- [49] Wittgenstein L. (1953) *Philosophical Investigations*, tr. by G. E. M. Anscombe. Prentice Hall, 1999
- [50] Ziemke T. (2003) What's that thing called embodiment? In: *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Lawrence Erlbaum, 2003 (<http://www.ida.his.se/~tom/cogsci03.pdf>)
- [51] Ziemke T. (2004) *Cybernetics and Embodied Cognition: On the Construction of Realities in Organisms and Robots*. *Kybernetes*, to appear (<http://www.ida.his.se/~tom/HvF.www.pdf>)