CHALLENGES FOR COGNITIVE SYSTEMS RESEARCH

A. Gomila & V. Müller

1. The task and how to approach it

a) Several ways to proceed -to be avoided

i) as Hilbert's formalist program for maths (more recently, DARPA's 23 challenges for maths): a list of problems, loosely related among them -it is possible to work in one disregarding the others

> an integrated approach to cognition is required

ii) standard way in AI-Robotics: "grand challenges" (like DARPA's driverless vehicles or RoboCup) -what matters in success, not how

> not necessarily an advance in understanding cognition, task-specific “tricks”

iii) as an internal agenda for a theoretical approach to cognition

> the proposed challenges may not be recognized as such by other approaches within the field

b) The goal:

i) provide a conceptual map of related issues, in a theory-independent way, that can provide orientation regarding what it is achieved, what's next, how issues relate to one another –and to do so providing milestones, scalable dimensions of progress, which are not bound to be dead alleys;

ii) do not restrict just to human-inspired, or human-like, artificial systems, even if human-like cognitive systems may be an outstanding goal, given the central interest in interaction between humans and artificial systems; but even in this latter case, no need to restrict to humanoid robots, but any form of cognitive interaction;

iii) not against the best practices/research programs on offer, but taking advantage of them to provide a common plan and vision, a consensus on what should be done first, and what counts as success.

2. What is a cognitive system -can there be artificial ones?

The way this question is answered, though, is critical to the specification of challenges.

At this basic starting point a critical split can be found in the field, between those that take for granted that cognition is computation, and those that, inspired by artificial life, establish a stronger connection between life and cognition, and view cognition as adaptation.

In order to avoid getting stacked at this starting point, we propose a definition of cognitive system that brackets the question of implementation and allows for diversity (rather than a rigid hierarchy of orders of complexity for cognitive systems): **a cognitive system is one that learns from experience and uses this knowledge in a flexible manner to carry out its goals.**

Notice the three elements in the definition: “learning from experience”, “flexible deployment of such knowledge”, and “autonomy” (own goals).

Of course, this is not an innocent or ecumenical notion -as a cursory attention to the debate on "minimal cognition" reveals. But it captures the central cases any approach to cognitive systems has to pay attention to (any definition involves its own borderline cases). Thus, it avoids considering all living beings as cognitive (reactive, reflex-like systems do not qualify). It allows for non-individual learning –or more precisely, it does not rule out evolution as a learning process at the supra-individual level-, but it emphasizes the connection between the learning experience and the flexible use of the knowledge (thus, adaptation does not guarantee flexible use of knowledge; thus, morphology by itself, even if it is the outcome of an evolutionary process of adaptation, does not qualify as knowledge). On the other hand, it requires more than computation for cognition: a meaning, a perspective, relative to one's goals.

The different aspects of this definition provide the ground to develop the challenges: learning from experience, knowledge, flexibility, own goals, in an integrated way.

3. Dealing with an uncertain world

Natural cognitive beings constitute a way to deal with an uncertain world (contraposed to the most common way: to adapt to just a robust subset of parameters in a rigid manner; cognitive systems exploit the information available in the environment to adapt). This suggests a relational understanding of world as what's relevant for the systems (as the old notion of "Um Welt" proposed): those parameters that may be relevant to our goals. By learning, cognitive systems try to discover the regularities, constancies, and contingencies, that are robust enough to provide guidance. Learning, though, should not be seen anymore as a passive recording of regularities, but as an active exploration (just like the role of infant play in motor development). In addition, given the relevance of relational contingencies, materials become important.

Talking of an "uncertain world" avoids the ambiguity involved in the notion of "impredictability", which can be applied both to the world and to the behavior of the system, but from an epistemic, rather than ontological, point of view. An uncertain world need not be a noisy one, just a complex one; on the other hand, a cognitive systems is one that behaves in ways that are not predictable from just the specified information about its structure, rules, or inputs.

-> Measures of progress:

- Ability to deal with increasing degrees of world uncertainty -but allowing for increasing variability (changing lighting conditions, distance, size,….); the jump from "virtual" to real environment; exploration of different environments (water, air, high temperature,…);

- Moving from “generational” ways of changing the system (as in genetic algorithms) to “tune up” with relevant kinds of information, to individual development of practical skills, of "ways to do things" in a proficient manner, through practice (rather than just habituation); both take time, but the time-scale is different: generational vs individual;

- In general, a motivated account of the "initial state" of the system, plus self-organizational development, is required –rather than “ad hoc” assumptions to make the system work;

- Exploration of the properties of new materials, sensors, actuators,... as a way to further explore the environmental “affordances”.

4. Learning from experience

This is probably the area where most efforts have been dedicated. There are a multiplicity of techniques and algorithms (broadly, the machine learning area) that try to account for this basic cognitive ability: statistical learning, supervised learning, reinforcement learning, Hebbian learning, dynamic context adaptation, explicit representation,… In general, all of them work with abstract data sets, rather than with real environments, and assume a passive view of the system (which is conceived as computational). This seems far from the way natural cognitive systems learn from experience: in an active, situated, way; by exploring the world; and by reconfiguring one's own skills and capabilities. On the other hand, “annotated” data sets can be seen as a form of social learning, but again passive rather than active.

-> Measure of progress:

a) Development of perceptual "representations" that can guide behavior (Bernstein's problem: combinatorial explosion of degrees of freedom --> via active exploration);

b) Finding structure in one's experience: the recognition of meaningful situations by active exploration –exploring the contingency between the data and the systems’s actions.

c) Finding analogies across domains; that is, relational simmilarities, rather than just superficial (sensory) simmilarities.

5. How to understand knowledge

Knowledge is the outcome of learning. The current challenge concerns the classical problem of knowledge representation. Classical AI got stacked with the idea of explicit, formal logic-like, representation, and reasoning as a kind of theorem-proving. Together with the aim to formalize expert (or common sense) knowledge, it could not solve the frame problem, the grounding problem, etc.. New approaches drive attention to practical, embodied, context-dependent, implicit, knowledge skills. But it is not clear yet how this new approach can be carried out: how knowledge is implemented, is stored (and how it is accessed, see next section). Unrealistic success of machine learning methods for classification tasks (via pattern recognition). Most promising approach: brain-inspired dynamical models –the knowledge in the topology of the network of processing units, plus its coupling to body and environment.

-> Measures of progress:

a) Advances from pattern recognition in the input data to relational laws (affordances) in the environment;

b) Advances in multisensory integration, rather than just sensor fusion; sensory-motor contingencies taking different sensors into account, for different environmental dimensions (visual –spatial, auditory –temporal,...), plus propioceptive information as disambiguating;

c) Brain-inspired networks of control (relative to each kind of brain, and each kind of body), for sensory-motor coordination in different tasks: need to go beyond navigation; in particular: a metrics for increasing the repertoire of behaviors available to the system;

d) Development, out of this basic, relational, understanding, of a detached, abstract, view of the world (objective knowledge). Psychology teaches that flexible knowledge requires some form of recoding, which is the key to abstraction, to making it adequate to novel, not exactly identical, situations. It can be said the neural networks (specially in their sophisticated forms) account for such abstract recoding, but it is doubtful; a different approach is to use layers of neural networks, where the higher level takes as inputs the patterns of the lower, sensory, layers, but up to now this is done "by hand". Still another approach, of Vygotskian inspiration, views in the use of public symbols the key to understand cognitive, abstract, re-coding.

6. Flexible use of knowledge

Extracting world regularities and contingencies would be useless unless such knowledge can guide future action in real-time in an uncertain environment. This may require in the end, as anticipated above, behavioral impredictibility, which is a property than runs contrary to the technical requirements of robustness and reliability for artificial systems (to guarantee safety, as the principal engineer's command). The critical issue for flexibility is related to how the knowledge is "stored" (see previous section), and therefore, how it is accessed. The major roadblock to carry this out –regardless of approach- is combinatoral explosion.

-> Measures of progress: Different strategies are actively explored as ways to reduce/constraint combinatorial explosion; it is not possible to establish a clear set of milestones at this point; need for exploration of new ideas (different programmes, but may be not incompatible: possible convergences):

a) simplifying the requirements for cognitive control (context-sensitivity of the "decision", distributed adaptive control architectures, potential conflicts adjudicated through accessibility, timing)

b) making the controller change in a stocastic way and select the variations that work better (genetic algorithms) –drawback: no individual learning, no flexible deployment of knowledge

c) dynamicist approaches, system criticalities in the state space, force-fields metaphors for distributed activation

d) exploiting the body (sensors and actuators) to constraint the options (morphological computation) –however, this raises a parallel problem of combinatorial explosion of “degrees of freedom” in the actuators;

e) advances in "schematization" of sensory-motor contingencies, but recoding (abstraction)

f) use of heuristics ("fast and frugal", non foul-proof, context-sensitive, procedures), instead of algorithms –but activation

g) emotions as quick valuations of situations, on simple hints –brain-inspired models of the reward system, of the amygdala; reinforcement-based expected reward, rather than calculation of expected utility

7. Autonomy

Autonomy is related to agency, and agency to own goals. It requires internal motivation, a sense of value. It also requires some kind of "self-monitoring": an internal grasp of one's cognitive activity is required to make possible the "internal error detection" (Bickhardt), as the central cognitive capacity of self-monitoring (involving both whether the behavior matches the relevant intention, and whether it is carried out as intended).

In systems like us, this property is achieved by a double control architecture: the autonomous nervous system (including the hormonal one), plus the central one, plus their links. In general, a cognitive system involves a basic regulatory system, that implicitly defines the needs and requirements, the motivations and homeostatic goals of the system, and which requires internal sensory feedback to keep the system within the range of vital parameters. In addition, a central system allows for more sophisticated forms of environmental coupling, for informational management, for memory and learning, and for control contingent on such previous experience. A full-blown agent, from this point of view, is one which is capable to generate new behavior appropiate to new circumstances (which seems impredictible just given the situation); it requires self-organization, a homeostatic relationship with the environment of self-sustained processes (very far from current technology). It may also require the ability to "work off-line", to recombine previous experiences, and to to the test of imagination the new options.

-> Measures of progress:

a) From programming of all possibilities (look and search strategy in the problem space) to let the artificial system "go beyond" the programmed, to modify itself, to choose among several options, to choose which knowledge to use,…

b) From systems with externally imposed goals in a non-previously specified manner in a previously specified environment (simplified, virtual); to similar systems able to deal with non-previously specified environments; from systems that can choose among several pre-specified goals according to circumstances, to systems that can develop new goals; from systems that can modify/change themselves according to circumstances to systems able to solve internal motivational conflicts (change goals) .

c) From "simple" systems, whose behavior depends upon a few parameters, to more "complex" ones –by increasing the number of parameters, and letting them to interact non-linearly, complexity follows.

8. Social cognitive systems

Social cognitive systems address this learning process in a facilitated way, by starting in a simplified, structured environment; by receiving feedback and scaffolding from others; by using others as models.

Of course, this creates a specific problem of social learning: to find out in the first place what parts of one's world are other cognitive systems, and to discover the regularities, constancies, and contingencies, that are relevant in this area.

It also opens up a promising area of interaction among cognitive beings, both natural and artificial.

-> Measures of progress:

a) Systems able to interact with other systems in increasingly complex ways - from simple synchronization, to imitation, to emulation, to cooperation, to joint action.

b) Systems able to develop "common worlds", a common understanding of how things go (share knowledge, distribute tasks according to abilities,…)

c) Increasing "mental" abilities -which cannot be dissociated from artificial systems abilities (to recognize "sadness" or "rage" in a human the artificial

d) Proficiency in linguistic interaction

As intended, progress in one challenge is not independent on progress on many others –the typical property of cognition is an integration of capabilities and elements. It is not possible, though, to establish milestones at this global level, because of the intrinsic diversity of cognitive beings. What it does seems adviceable at this point, though, is to emphasize integrated systems over specialized algorithms. Classical AI has worked under the assumption of modularity, as engineering in general: the goal is to add new facilities to a system without having to change it. There is reason to doubt this assumption is going to work for cognitive systems –the scale-up problem is a really one. New capabilities may require some sort of reorganization, in non-principled ways. Hence, a final, global, challenge, concerns this problem of scaling-up cognitive systems –which may call for an evolutionary approach.