

EUCogII on “Challenges for Cognitive Systems”

Vincent C. Müller
Anatolia College/ACT

EUCOGII

Overview

1. Introduction on Workshop
2. Activities Carried Out/Working Groups
3. Systematic Results

I - Introduction

- Rapperswil, 28-30.1.2011
- 37 participants were selected from about twice as many applicants - essentially on the basis of seniority and a good mix.
- online questionnaire, comment on proposals and to propose urgent issues
- Starting point:
“The set of challenges should:
 - provide a long-term vision and fruitful orientation for present work
 - be theory and strategy neutral (not fashion dependent, open to new approaches)
 - not be domain specific, not be oriented towards toy problems or scenarios
 - be systematic
 - be measurable”

II a Narrative

- Systemic challenges ordered by prominence in questionnaire results:
 1. learning and development
 2. perception (and integration)
 3. categorization, representation and inference
 4. action and goal selection
 5. communication and cooperation with other agents
 6. distributed cognition and culture
 7. experience (conscious or other)
 8. motor coordination

- Benchmark challenges
 1. Evolution & development
 2. Action selection
 3. Categorization, representation & inference
 4. Cooperation & Communication
 5. Embodiment
 6. “The whole iguana”

II b Proposals from Working Groups

- 1. Evolution & development
Learn an arbitrary task within the same learning environment and to the same level as a biological creature. The biological creature will be drawn from an increasingly complex set of biological creatures ...
- 2. Action selection
An artificial cognitive system enters an unknown, possibly crowded supermarket, with a list of items to be bought for a human. (Multiple levels of action selection, multiple goals.) - metrics: complexity of environment, speed & quality of success.
- 3. Categorization, representation & inference
Place objects found in a room in the correct boxes on the table. - metrics: number of objects, complexity of categories, distractors, difficulty of placing objects

- 4. Cooperation & Communication

Cloud of humans, robots and machines - with distributed sensing, cognition, actuation, and communication across humans and machines, who give a small part of their resources in order to help towards a task; distributed and dynamically adapting to change of available resources. - metrics: user satisfaction & participation, success at given tasks

- 5. Embodiment

a) Play tennis - metrics: performance, energy consumption, program complexity, price

b) Human robot companion - metrics: trustworthiness, energy consumption, program complexity, price

- 6. “The whole iguana”

Benchmarking the success of a full cognitive agent within an environment. Since both degrees of success and complexity of environment are relevant factors, this can be seen as a two dimensional space with “environmental complexity” on one axis and “agent coping ability” on the other.

III Systematic Results

A) We need systematically motivated benchmark challenges

- The tension between the two aims is inherent in “a long-term vision and fruitful orientation for present work”.
- We cannot expect to formulate benchmarks once and for all. (This is not mathematics; Hilbert’s 23 Problems.)

III Systematic Results

B) Benchmarks involve a) measurable success or b) measurable variation in complexity of the environment

- a) speed, quantifiable output, comparison to other agents (natural or artificial) or 'quality' of output (+ use of resources)
- b) enumerate relevant factors or use probabilistic measures.
- Ability to establish clear comparable metrics is inversely proportional to the degree of achievement.
- Real environments can only be specified to a degree, i.e. they cannot be formal.

III Systematic Results

C) Benchmark challenges must test an entire autonomous system in an environment

- System performance in particular abilities is strongly dependent on overall features of the system, involving a host of different abilities - even if the benchmarks measures one.

D) Benchmark challenges must specify 'cheating'

- If benchmark challenges are set with respect to success in an environment, we a) ignore internal workings and b) allow any working solution.

Summary

- A) We need systematically motivated benchmark challenges
- B) Benchmarks involve a) measurable success or b) measurable variation in complexity of the environment
- C) Benchmark challenges must test an entire autonomous system in an environment
- D) Benchmark challenges must specify 'cheating'