LUND UNIVERSITY

Robot Design
Concluding Lecture

Jacek Malec
Department of Computer Science
Lund University

July 23, 2007

AI@CS
Department of Computing Science

LUND UNIVERSITY

Outline

- Reflections
- Lessons to learn
- Where to go from here
  - Specification, software engineering
  - Embedded system design
  - Adaptive Control
  - Formal methods
  - Agent specification

AI@CS
Department of Computing Science

LUND UNIVERSITY

Reflections to make after the contest:

- What went wrong?
- What was the reason?
  - external explanation vs. design-centered
  - half-empty-bottle effect
- How could we have changed the outcome?
  - Morphology?
  - Feedback?
  - Algorithms?

AI@CS
Department of Computing Science

LUND UNIVERSITY

Lessons to learn

- About the design process itself:
  - Design is a tedious, repetitive process: very seldom the first idea is (completely) right
  - Design is a process requiring possibility of testing/verifying the ideas
  - Successful design requires deep reflection over its outcomes
- About complexity:
  - It is defined by possible interactions between the robot and its environment
  - Needs to be taken into account
- One should attack a problem in a principled way.

AI@CS
Department of Computing Science

LUND UNIVERSITY

Where to go from here?

- Specification, software engineering
- Embedded system design
- Formal methods
- Agent specification
- Robotics
- Control theory, adaptive control
- ...

AI@CS
Department of Computing Science

LUND UNIVERSITY

Robot applications

- Search and rescue
- Fire fighting
- De-mining
- Medical technology
- Production technology
- Exploration: ocean, space
- Colonisation: ocean, space
- Entertainment
- Defence

AI@CS
Department of Computing Science

Ethics

[Kale Harbick, USC]

- Knowledge increasing faster than wisdom
- How much intelligence should robots be given?
- Where do we draw the line?

A rational agent

[Wooldridge, 2000]

An agent is said to be *rational* if it chooses to perform actions that are in its own best interests, given the beliefs it has about the world.

Properties of rational agents:

- autonomy (they decide);
- proactiveness (they try to achieve their goals);
- reactivity (they react to changes in the environment);
- social ability (they negotiate and cooperate with other agents).

Logical characterisation of agents

- How can we characterize rationality?
  - crucial terms: *beliefs* and *best interests*;
  - decision theory, game theory;
  - **logic**;
  - epistemic logics;
  - temporal logics, dynamic logics;
  - BDI logics.
- How can we build rational agents?
  - software architectures;
  - executable specifications;
  - **logic**;
  - BDI logics.

The omniscience problem

All normal modal logics obey the following:

- K axiom:

$$\models \Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$$

- necessitation rule:

$$\frac{\alpha}{\Box\alpha}$$

Consequences?

- infinite set of beliefs
- consistency of beliefs
- equivalence of beliefs

Ways of dealing with omniscience

We need to model:

1. bounded resources of agents;
2. incomplete reasoning mechanisms.

- Weakening the system - non-standard semantics
- Distinguishing explicit and implicit knowledge
  The two ways above suffer from partial logical omniscience!
- Removing closure properties
- Syntactification