Selecting Actions and Making Decisions: Lessons from AI Planning

Héctor Geffner
ICREA and Universitat Pompeu Fabra
Barcelona, Spain

Workshop on Modeling Natural Action Selection
Edinburgh, 7/05
Motivation

How are simple problems such as this solved by people?

- Work on the psychology has focused on problems that are hard for people (puzzles), yet . . .

- Even simple problems are computationally hard for a general problem solver if it does not recognize and exploit structure
A general problem solver must recognize and exploit structure in problems, otherwise computational complexity overwhelming.

In last 10 years, work in AI Planning and Problem Solving has produced robust techniques for recognizing and exploiting structure that have been evaluated empirically.

These techniques let a general problem solver adapt to the task at hand, and likely to be relevant for understanding how people find solutions to problems.
Techniques

Some techniques for recognizing and exploiting structure in problems that proved robust experimentally are:

- automatic extraction of **heuristic functions** from problems descriptions for guiding the search (heuristic function estimate cost to goal)

- **tractable inference** for reducing the search, eliminating it completely in many cases

- automatic **transformation of representations** so that certain hard inferences become computationally easy (**knowledge compilation**)
Example: Automatic Derivation of Heuristic Functions

- Assume a set of actions $a$ characterized by preconditions, positive effects and negative effects, and costs

- Computing optimal costs $g^*(p, s)$ for achieving arbitrary atom $p$ from state $s$ intractable, yet can be efficiently approximated as:

$$g(p; s) \overset{\text{def}}{=} \begin{cases} 0 & \text{if } p \text{ holds in } s, \text{ else} \\ \min_{a: p \in \text{add}(a)} [\text{cost}(a) + g(\text{pre}(a); s)] & \end{cases}$$

where $g(C; s) \overset{\text{def}}{=} \sum_{r \in C} g(r; s)$ when $C$ is a set of atoms

- Distance to Goal from state $s$ can then be approximated by heuristic function

$$h(s) \overset{\text{def}}{=} g(\text{Goal}; s)$$

and used for selecting actions; e.g., pick action that takes you closest to the goal.

Issues: Domain-generality vs domain-specificity

• Domain-general mechanisms questioned by evolutionary psychologists and cognitive scientists from the fast and frugal heuristics school

• Yet on the one hand, domain-specificity brings own problems: how many domains, what are the borders, how modules selected, . . .

• On the other hand, the recent work in AI shows that general and adapted not necessarily in conflict; key is recognition and exploitation of structure

• E.g., heuristics above are fast and frugal (i.e., linear-time) but also general; their form resulting from the actions in the domain

• There is no question, however, that key features built-in by evolution in the DNA (E. Baum 1994)
• Solutions of many models, such as those involving *uncertainty* and *feedback*, are *functions* (*policies*) mapping *states* into *actions*

• These functions can be *represented* in many ways (e.g., as condition-action rules, value functions, etc), and can be *obtained* in many ways as well; e.g, *policies* can be
  – *computed automatically* from problem representations in *Al Planning*
  – *written-by-hand* in suitable architecture in *Behavior-based Al*
  – *hardwired-in-brains* by process of evolution in *Behavioral Ecology*

• Representing and executing solutions, however, while challenging, is different than coming up with the solutions in the first place which is what Al Planning is about.

• Whether this is a requirement of intelligent behavior in animals is not clear although it seems to be a distinctive feature of intelligent behavior in humans.
Emotions

- Emotions no longer viewed as obstacle for good decision making, but rather as aid (Damasio 1994):
  
  “let emotions be our guide” (Ketelaar and Todd 2001)
  “emotions help humans solve the search problem” (D Evans 2002)

- Emotions apparently summarize vast amounts of information (beliefs, preferences, costs, etc).

- The key computational question is how emotions accomplish these appraisals in real-time.

- AI can help here as well; e.g.,
  
  – Work on theory compilation (Darwiche 1990) suggests how similar appraisals can be done in linear-time over compiled representation; while
  – Work on the automatic extraction of heuristics suggests how numbers approximating cost information can be computed in linear-time as well
Summary

• Balancing generality and efficiency is a key concern in agent design

• Both goals attainable if structure of problems recognized and exploited

• Recent work in AI shows this is possible and how:
  – automatic extraction of heuristics for guiding search
  – tractable inference for eliminating search in many cases,
  – theory compilation for speeding up inferences

• Ideas underlying these techniques likely to be relevant for understanding human problem solving, and computational basis of emotions

• Exploitation of structure also central in E Baum’s *What is Thought*, MIT Press 2004, but in context of evolution; both views however are complementary