Contents

1 Problems of deliberative architectures

- deliberation requires **accurate world models**
- hard to offer real-time guarantees on performance:
- solving any given problem takes longer than an **equivalent reactive implementation**
- solving different problems takes different amounts of time

1.1 Deliberation is inherently serial

- generation of each alternative is **inherently serial**, since each step relies on the state **produced by the previous step**
- the number of courses of action grows exponentially with the length of the solution, rather than linearly in the number of percept-action pairs
- the number of alternatives an agent can pursue in parallel is bounded by the number of **processing units available**
- consideration of some alternatives **must be deferred**, i.e., processed serially

1.2 Deliberation takes more time

- in a reactive architecture, if we know which action(s) to perform in a given situation we can perform them **immediately**
- all other things being equal, **deliberation will take more steps** than simply reacting, since we have to (serially) generate and compare alternatives
- if the agent doesnt learn, it **must re-solve a problem each time** it encounters it, and will always be slower than a reactive agent
- in a dynamic environment a deliberative agent may take **too long to select an action**, e.g., if the situation changes before deliberation is complete

1.3 Deliberation takes unpredictable time

- there is no one best problem-solving method (No Free Lunch Theorem)
- different problems are more or less difficult **for different techniques**
- it is difficult to tell how hard a problem is **just by looking at it**: we have to try and solve it
- we dont know how long it will take to find a solution to a problem **in advance**
- time required by a deliberative agent **will vary** from problem instance to problem instance

1.3.1 Approaches to deliberation time

- **ignore it** and hope for the best, e.g., classical planning
- make deliberation run faster, e.g., by using non-optimal algorithms which **sacrifice solution quality for speed**
- try to **predict** how long deliberation will take to solve the current problem, e.g, empirical AI
- **adapt** the amount of deliberation performed **to the time available**, e.g., anytime algorithms
- accept that deliberation may occasionally **be too slow** for some environments or some problems, e.g., monitor the environment and replan if the environment changes

2 Accurate world models

Agent must be able to:

- **update its model** of the current environment based on its percepts
- generate the **counterfactual part of the model** by correctly predicting how the world will change, e.g., as a result of its actions

if the world model is incorrect, correct deliberation may **select the wrong course of action**

2.1 Representing the real part

it is difficult to create and maintain the veridical model of the current state of the world required for deliberation:

- the information available to the agent may be **incomplete**, e.g., if the environment is only partially observable
- the information available to the agent may be **inaccurate**, e.g., if the environment has changed since the representation was last updated or the agents sensors give incorrect information

2.2 Representing the counterfactual part

To generate hypothetical future states which are useful, the agent must be able to accurately predict how the world will change.

- as a result of the **agents actions**: agents actions are often assumed to be infallible;
- as a result of **exogenous changes**: actions of other agents or environmental changes are assumed to be predictable

2.3 Approaches to errors in representation

- **ignore them** and hope for the best, e.g., classical planning
- try harder to make the representation **fit reality**, e.g., by using more and better sensors, or by trying to ensure that actions really are infallible, e.g., by implementing them as robust reactive behaviours
- explicitly **represent and reason about uncertainty**, e.g., probabilistic representations, decision theoretic approaches
- **interleave planning and acting**: plan only a little way ahead and keep checking that things are going according to plan; if not, replan.

3 Empirical AI

- difficult to tell how hard a problem is for a given problem-solving technique **without trying to solve it**
- empirical AI tries to characterise which problems are hard or easy for a given technique
- most of this work has been in the areas of optimisation, planning and constraint satisfaction problems
- concerned with which problems are soluble at all

4 Anytime algorithms

- interruptible algorithm which produces a sequence of solutions of monotonically increasing quality
- anytime algorithms can be used in two ways:
  - to compute the optimum amount of time to spend deliberating to **maximise the utility of the outcome**, e.g., decision theoretic approaches
  - **deliberate for as long as possible** before acting, on the assumption that deliberating for longer will produce a better solution

5 Planning in dynamic environments

Reference section

placeholder