Seminar 5 - Climate policy under uncertainty, optimality, policies

Adrian Perez Keilty

Chalmers University of Technology, Göteborg, Sweden

Exercise 5.1:

Not every decision can be applied in every state. Decisions (controls) that can be applied in a given state are said to be feasible for that state. Give an example of a simple control problem in which certain controls are not feasible. What could be the type of a feasible predicate?

Nicholas provided a nice example of this. Consider the scenario where a machine is exploring a two dimensional maze. It can go either left, right, straight or backwards. But clearly, not every direction can be applied in every position, e.g, if the machine is facing a wall, it should not be able to move straight. The type for a feasible predicate should be Maybe.

Exercise 5.2:

Even a two-step decision plan could be unfeasible. Explain why.

This exposes the issue with unexpected events taking place in between the two steps that have been planned. For example, when planning to walk from point A to point B an unexpected obstacle may appear after we depart from A and before we arrive to B.

Exercise 5.3:

Under stochastic uncertainty, it is generally not a good idea to take decisions which are optimal for expected states. Explain why. Give an example in which this is in fact the worst that one can do!

I will mention Nicola's example on the football goalkeeper, who according to the stochastic data, decides to stay in the middle of the goal instead of guessing a direction to jump to (left or right)

Exercise 5.4:

Explain the at most $1 + n \cdot |Y|$ *estimate.*

One optimal policy is reached after trying out all possible decisions ($|Y|$ total decisions) for every possible state ($|X|$ total states), i.e, $|Y|^{|X|}$. Hence, taking n optimal policies will take at most $n \cdot |Y|^{|X|}$. Since taking $1+n$ optimal decisions under uncertainty requires computing n optimal policies, this means at most $n \cdot |Y|^{|X|}$ optimal decisions need to be computed. (Not sure where the + 1 come from?)

Exercise 5.5 (a generation dilemma):

Should a generation in GU do a or b? The answer is: it depends. Explain on what it might depend.

With this model, only *time* seems to be the determining factor. If the BT phase does not extend for an unreasonable amount of time, then b is the way to go, otherwise it might be safer to choose a.

The other determining factor is the probability of staying in GU if a is chosen. Evidently, if this probability is low, then it would objectively be better to choose b and transition through BT instead of ending up in B permanently with a non negligible probability.

Exercise 5.6 (a generation dilemma):

Put consistent probabilities on the edges of the transition graph above.

A consistent choice for probabilities to estimate the likelihood of having phase GU or B as outcomes would be

$$P(\{\text{“decision a leads back to GU”}\}) = 1 - \alpha$$

and

$$P(\{\text{“decision a leads back to B”}\}) = \alpha$$

for some close to negligible value α (decision b leads to deterministic outcome).

Exercise 5.7:

Try to formalize the cartoon of step 1 (Sequential decision problem) (abstract away the details of specific decision problems) in Agda

We can define the possible states (X t) at a given step t : Nat and their corresponding possible options (Y t x) as well as the evolution of the system through the transition (next) function. Rewards can be modeled as values of a special type Val computed in terms of the decision step t, the current state x : X t, the decision made y : Y t x and the resulting next state:

```
-- Set of states
postulate X : Nat → Set

-- Set of controls
postulate Y : (t : Nat) → X t → Set

-- M-structure pf possible next states
postulate next : (t : Nat) → (x : X t) → Y t x → M (X (suc t))

-- Rewards
postulate reward : (t : Nat) → (x : X t) → Y t x → X (suc t) → Val
```