Verified dynamic programming

Nicola Botta[1,2]

[1]Potsdam Institute for Climate Impact Research
[2]Chalmers University of Technology, Gothenburg, Sweden

FP winter meeting, Jan. 2020

Outline

- Bellman equations
- Where do they come from? A sketch of DP
- Monads and extensional equality preservation
- Naive DP theories

Bellman equations

- Wikipedia, Bellman equation in Markov decision processes:

$$V^\pi(x) = R(x, \pi(x)) + \gamma \sum P(x'|x, \pi(x)) * V^\pi(x')$$

value of π when x' starting in x !

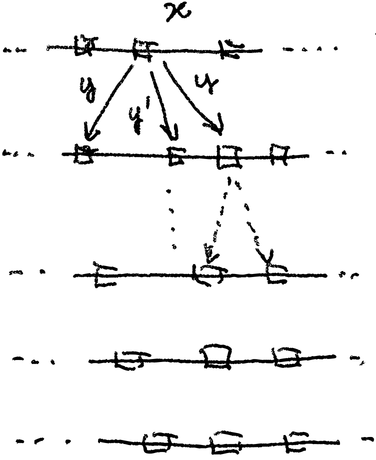
- Bertsekas' lecture slides on dynamic programming:
  - System $x_{k+1} = f_k(x_k, u_k, w_k), k = 0, ..., N-1$.
  - Control constraints $u_k \in U_k(x_k)$.
  - Probability distribution $P_k(\cdot|x_k, u_k)$ of w_k.
  - Policies $\pi = \mu_0, \ldots, \mu_{N-1}$, where μ_k maps states x_k into controls $u_k = \mu_k(x_k)$ such that $\mu_k(x_k) \in U_k(x_k)$ for all x_k.
  - Expected cost of π starting at x_0 is

$$J_\pi(x_0) = E\left\{g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(x_k), w_k)\right\}$$

value of π when starting in x_0 !

- Many more formulations ...

Where do they come from? A sketch of DP



$X_0\,,\ p_0 : (x : X_0) \to Y_0\, x$

$X_1\,,\ p_1 : (x : X_1) \to Y_1\, x$

$X_{n-1}\,,\ p_{n-1} : (x : X_{n-1}) \to Y_{n-1}\, x$

$X_n\,,\ p_n : (x : X_n) \to Y_n\, x$

X_{n+1}

$$V\ [p_0, p_1, \ldots p_n]\ x = R\ x\ y + V\ [p_1 \ldots p_n]\ x'$$

<u>where</u> $y = p_0\, x\,,\ x' = \mathrm{next}\ x\ y$

Monads and extensional equality preservation

$$ExtEq : \{A, B : Type\} \to (f, g : A \to B) \to Type$$
$$ExtEq\ f\ g = (a : Domain\ f) \to f\ a = g\ a$$

$$M : Type \to Type$$ (List, SimpleProb, Maybe ..)

$$map : \{A, B : Type\} \to (A \to B) \to M\ A \to M\ B$$

$$mapPresExtEq : \{A, B : Type\} \to (f, g : A \to B) \to$$
$$ExtEq\ f\ g \to ExtEq\ (map\ f)\ (map\ g)$$

Naive DP theories