Modeling and optimizing patient-management processes with POMDPs

Milos Hauskrecht

Computer Science Department, University of Pittsburgh

milos@pitt.edu

Decision making in medicine

Medicine:

- One of the most important areas of applied decision-analysis and decision-making

Typical methods of analysis:

- decision trees **vs.** MDPs and variants

OR and statistics communities are very active in health care applications:

Margaret L. Brandeau, François Sainfort, William P. Pierskalla (eds.) Operations Research and Health Care: A Handbook of Methods and Applications, 2004.

Medical therapy planning

Assume we want to model the decision making process of a physician for managing the patient

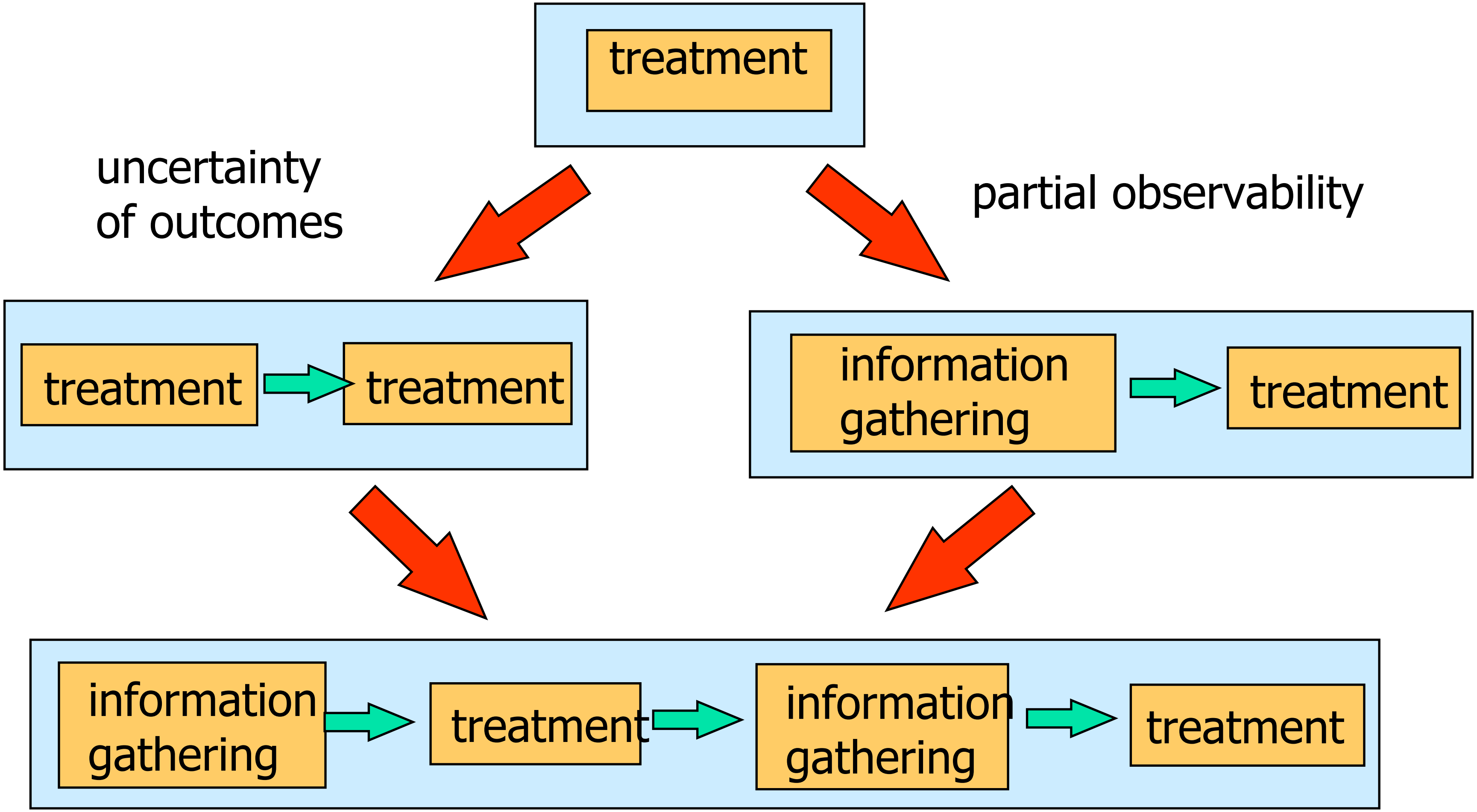
The model should represent:

- **action-outcome uncertainty**
  outcome of a therapy, surgery is uncertain
- **partial observability**
  underlying disease is not known with certainty
- **Complex temporal cost/benefit trade-offs**
  in between treatment and investigative procedures

Examples:

- management of chronic diseases (ischemic heart disease)
- management of a patient in the ER (acute chest pain)

Effects of uncertainty



Medical therapy planning

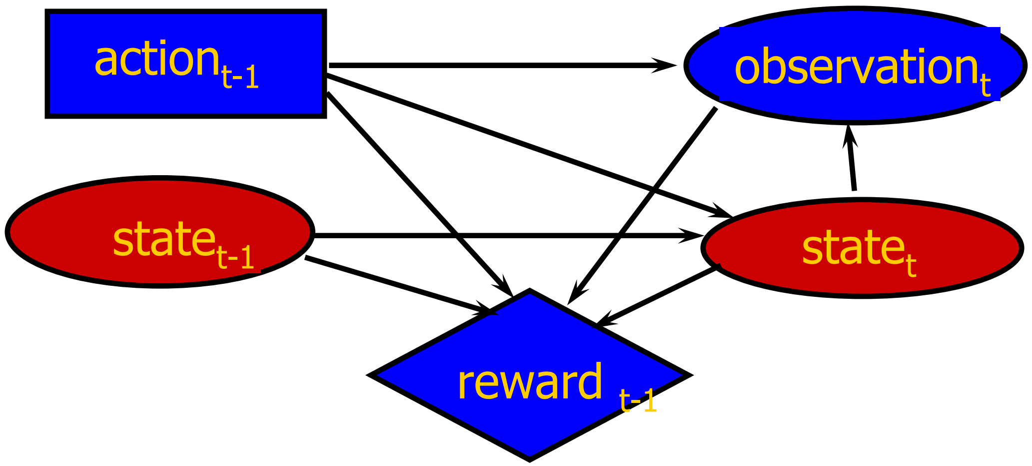
To identify optimal or near-optimal management actions we need:

1. **A model that represents**
   - **the dynamics of a patient state** under different interventions
   - **preferences of a patient/physician** combining patient state outcomes and action costs (in term of patient suffering)

2. **Computational methods**
   - For finding a policy or a decision for managing the patient with the optimal or near optimal cost-benefit tradeoff

Our solution is based on the POMDP framework

Partially observable Markov decision process



POMDP

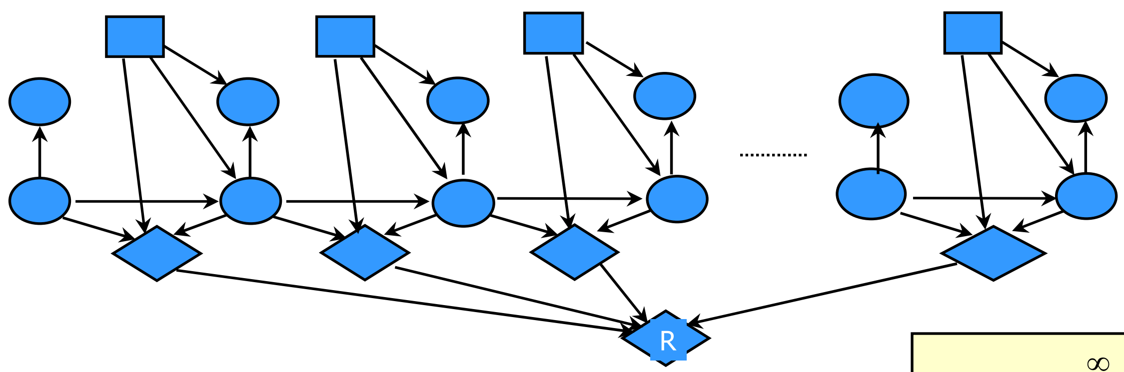
set of process states: S	diseases, disease stages
set of actions: A	treatment, investigative procedures
set of observations: Θ	test results, symptoms
transition model: $P(s_t \| s_{t-1}, a)$	disease dynamics
observation model: $P(o_t \| s_t, a)$	disease-observation relations
reward (cost) model $R(s_{t-1}, a, s_t, o_t)$	payoffs associated with a transition and new observation

Valuation model

Goal: assess the goodness of all possible dynamic behaviors resulting from following a policy π

Valuation model (criterion):

- combines future rewards over multiple steps
- expectation of outcomes for multiple possible behaviors



Infinite horizon, discounted model

$$E\left(\sum_{t=0}^{\infty} \gamma^t r_t\right)$$

Planning: find the policy optimizing the valuation model

POMDPs for medical planning

This talk:

- Problem of management of patients with ischemic heart disease (IHD)

Contributions:

- A factored POMDP with hidden and observable state components for IHD (Hauskrecht & Fraser AMIA 98, AIMJ 2000)
- factored value function approximation methods for solving this factored POMDP (Hauskrecht JAIR 2000, Hauskrecht & Fraser AIMJ 2000)

Management of patients with chronic ischemic heart disease

Ischemic heart disease (IHD):

- Impairment between heart oxygen supply and demand, usually due to the coronary artery disease

Main goal:

- optimal management of the patient with the disease

Management objectives:

- maximize the length and quality of life, minimize pain and suffering, economic cost o procedures

This work:

- Focus on a long-term management problem

A factored POMDP model for the IHD

State: factored state defined in terms of state variables

- Patient state variables: Coronary artery disease, ischemia level, chest pain, rest EKG, etc.
- Some variables are observed (e.g. rest EKG, chest pain) others are hidden (e.g. coronary artery disease)

Actions: 6 actions

- Investigative and treatment options

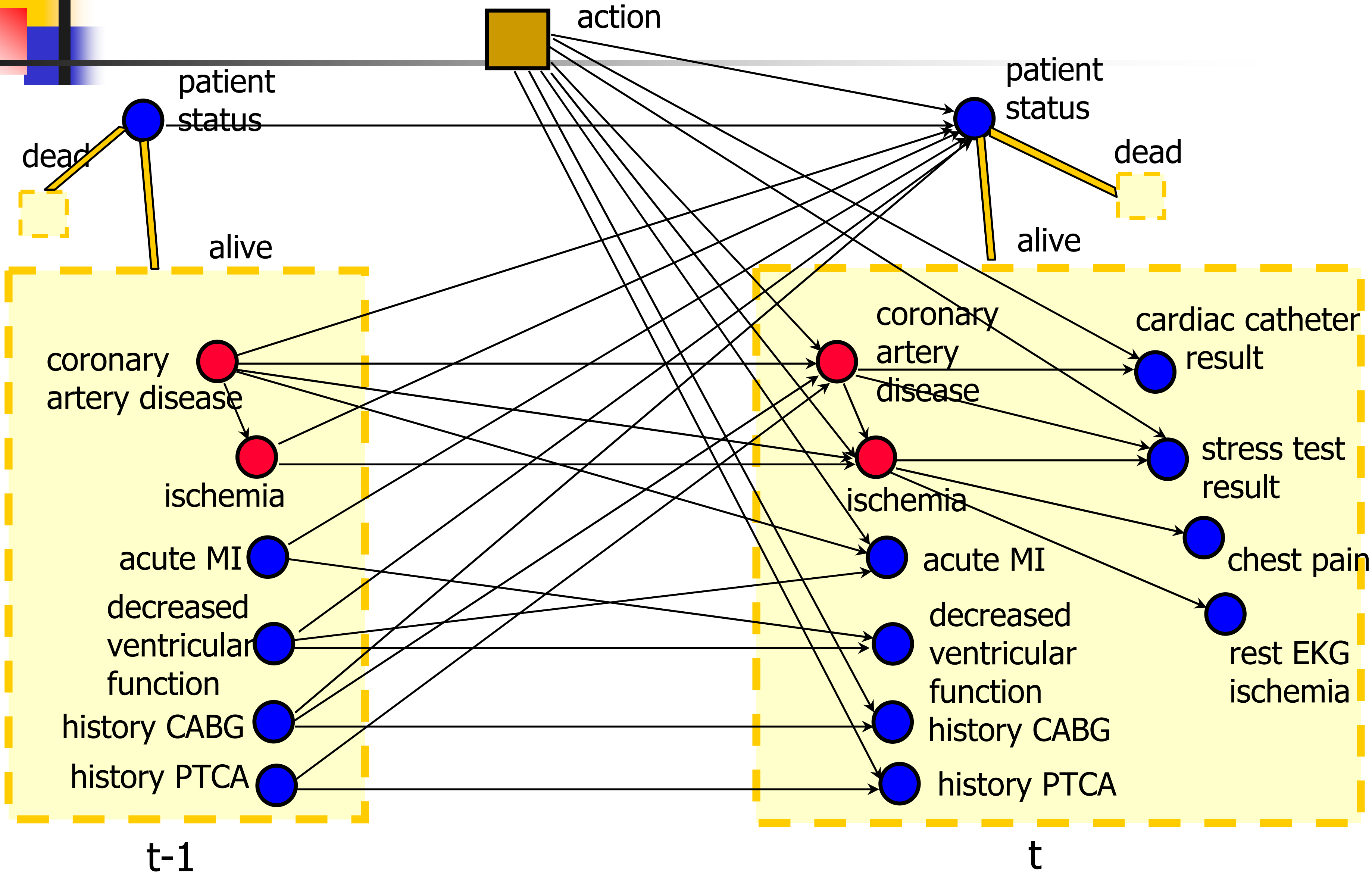
investigative actions:	treatment actions:
stress-test	**no-action**
angiogram	**medication**
	PTCA (angioplasty & stent)
	CABG (coronary bypass)

A factored POMDP model for the IHD

Transition and observation models

- **hierarchical dynamic belief network** with additional independence structure
- It models dependencies (independencies) that hold among state components in two consecutive time steps
- Observations and process state variables are treated the same
- Parameters of the model based on (Wong 92) and the estimates of a cardiologist

Dynamic belief network for the IHD



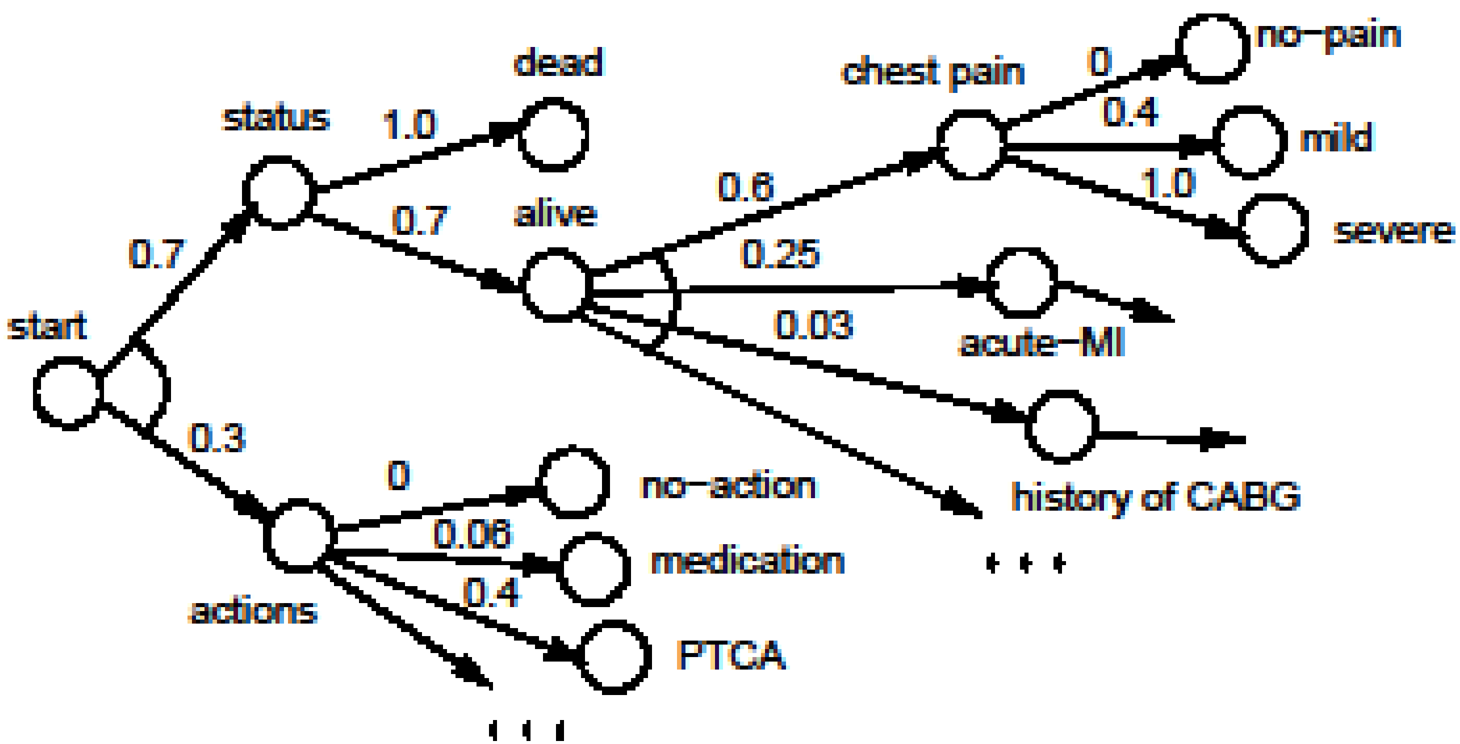
Legend: observable variables hidden variables action

Cost (reward) model for the IHD

Cost are associated with:

- Next states
- Actions performed

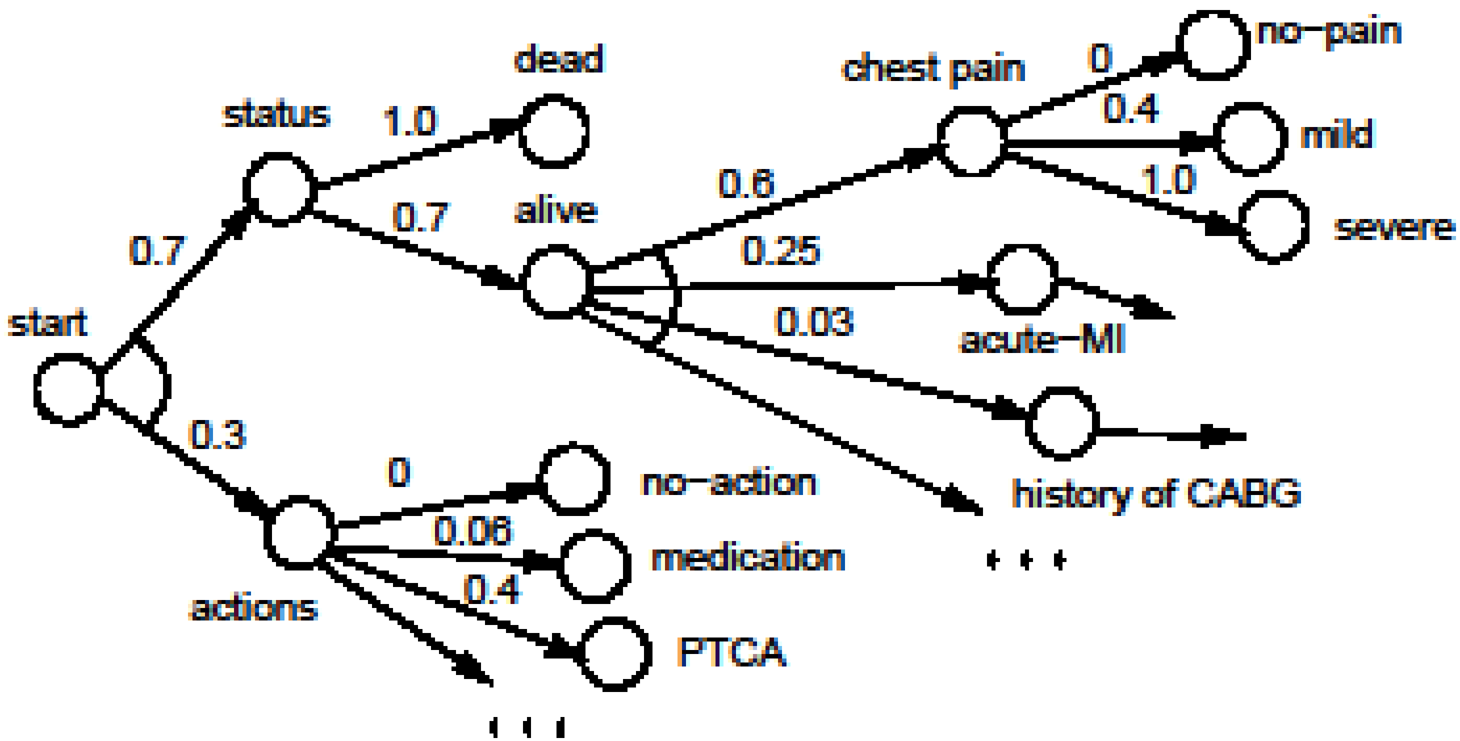
Cost model: acquired from the expert



Cost (reward) model for the IHD

Cost model: acquired from the expert

The **cost tree** represents relative importance of: actions and the next state; (next) state variables; and their values for defining the transition costs

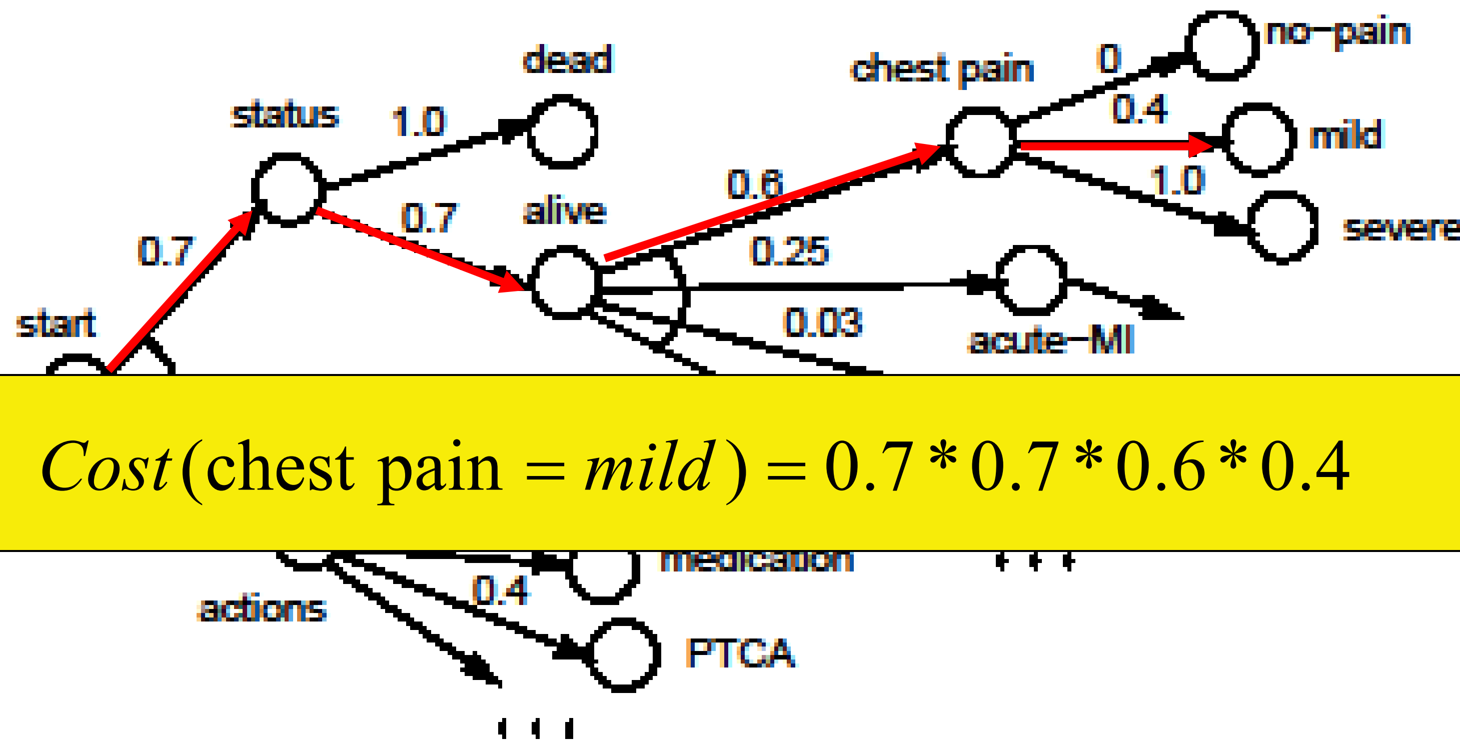


Cost (reward) model for the IHD

The cost model is factored:

$$Cost(s,a) = Cost(a) + \sum_{s_i} Cost(s_i)$$

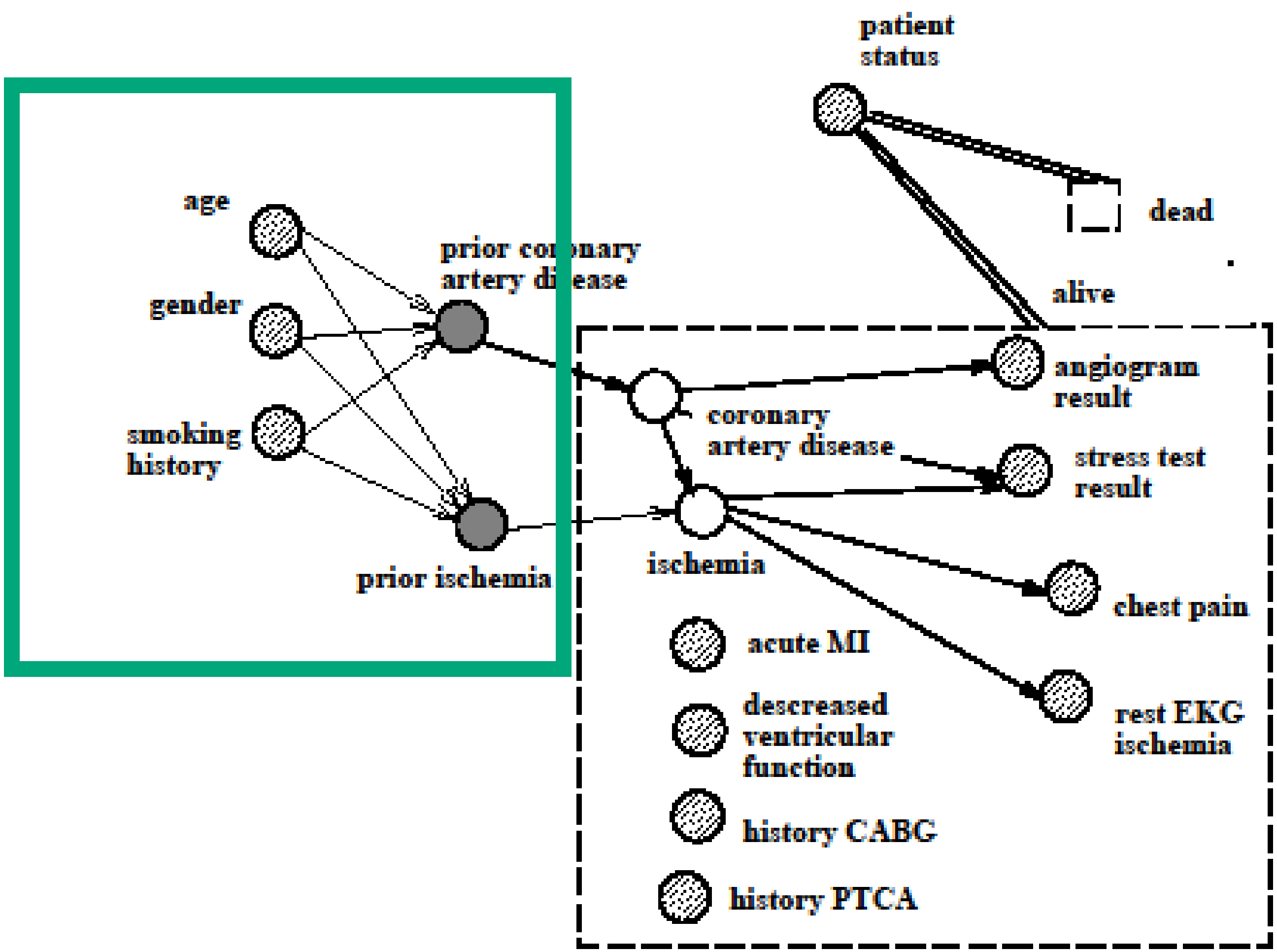
Individual costs are obtained from the cost tree, e.g.



$$Cost(\text{chest pain} = mild) = 0.7 * 0.7 * 0.6 * 0.4$$

A POMDP model of the IHD

Prior model: defines the initial belief state



Solving the DBN-POMDP for IHD

Exploitation of the DBN-POMDP structure

- combination of fully observable and hidden variables
  - ⇨ decomposition of the belief state into observable and hidden parts
  - ⇨ decomposition of the value function to a set of value functions over smaller belief space

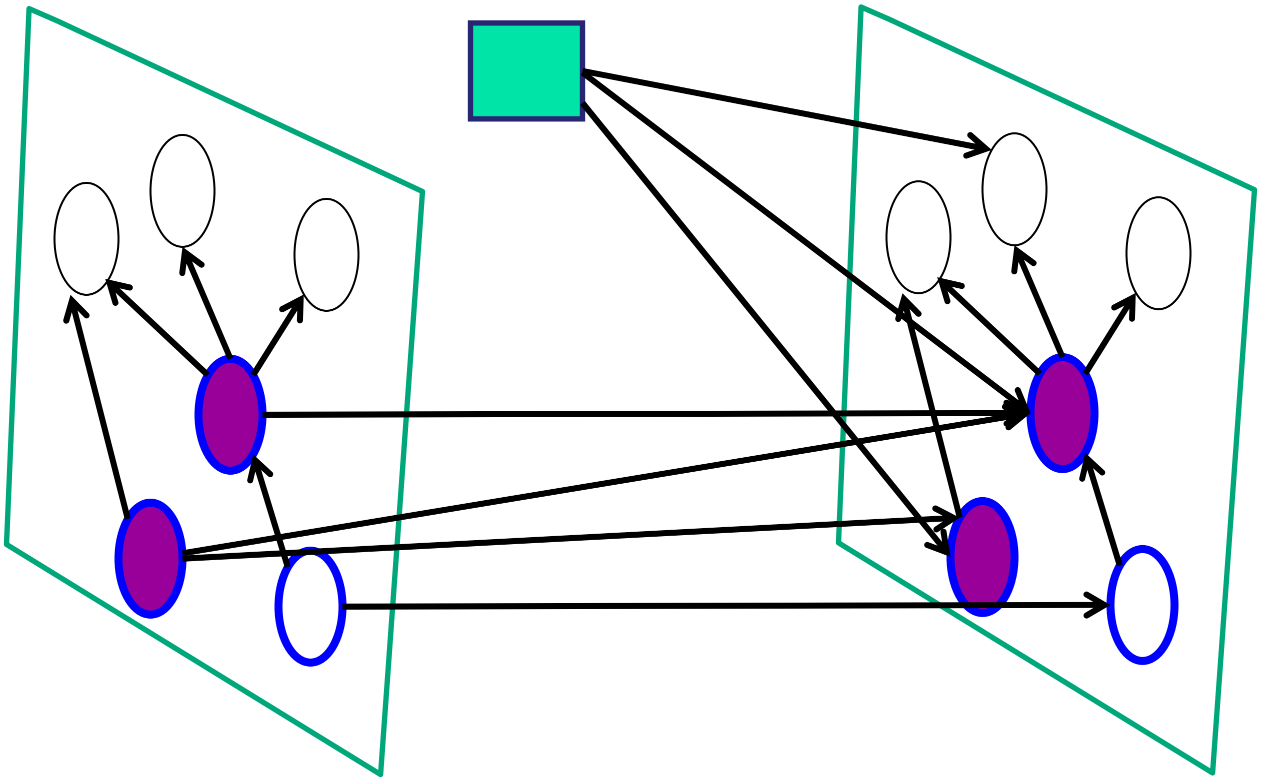
combined with value-function approximations

- Structured versions of the value function algorithms from (Hauskrecht AAAI97, JAIR 2000)
  Examples of methods implemented:
  - fast informed bound method,
  - grid (point) based linear function methods

DBN-POMDP

A state of the DBN-POMDP is defined by:

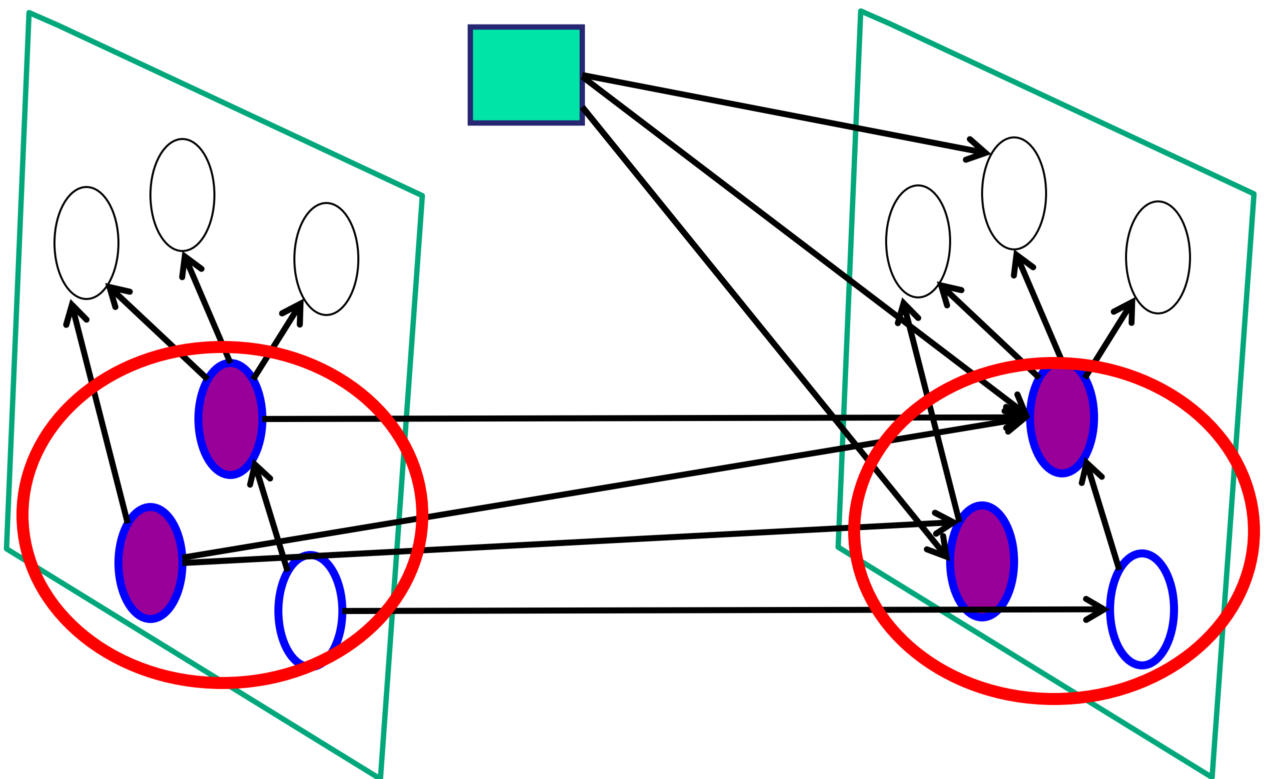
- A mix of observable and hidden variables and their values



DBN-POMDP

A state of the DBN-POMDP is defined by:

- A mix of observable and hidden variables and their values
- Note that a smaller subset of **process state variables** is sufficient to define the dynamics of the Markov process



DBN-POMDP

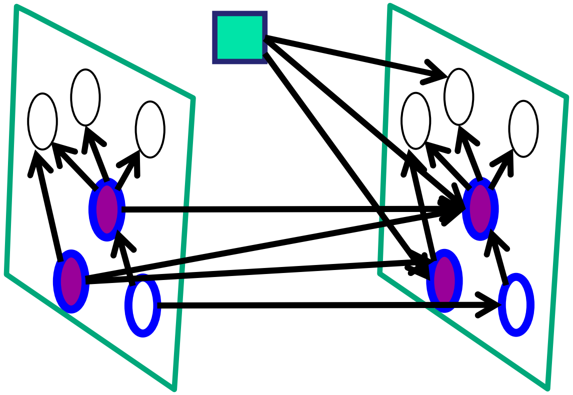
A process state for a DBN-POMDP:

- **Process state** (s,h) – a vector of values for observable and hidden process state variables
- **Observations** o - a vector of values for all observable variables.
- **Observable process state s**
  - A vector of values for all observable process state variables
  - Obtained from o by projecting (choosing) process state variable values, that is $s=proj_s(o)$

DBN-POMDP

Information state I for the DBN-POMDP:

- Restores the Markov property of the process
- $I = (s, b(.|o))$
  - Consists of the vector of values for observed state variable values + belief over values of hidden state variables



Solving DBN-POMDP

- **Bellman equation:**

$$V^{i}(I) = \max_{a}\left[R(I,a) + \gamma \sum_{o} P(o \mid I,a) V^{i-1}(\tau(I,a,o)) \right]$$

- **Information state update:**

$$I' = (s', b'(. \mid o)) = \tau(I,a,o) = (proj_{s}(o), \tau_{h}(I,a,o))$$

$$\tau_{h}(I,a,o) = P(h' \mid a,o,I)$$

- **Value function: a function over the belief state for each observable process state vector s**
  - **A pwlc representation of the value function**

$$V^{i}(I) = V^{i}(s, b(. \mid o)) = \max_{\alpha \in \Gamma_{i}(s)} \sum_{h} \alpha(h) b(h \mid s)$$

Value-function (VF) approximations

Value function approximations

$$V^*(b) \approx \hat{V}(b)$$

$\hat{V}(b)$:
- a function of simpler complexity
- computable efficiently

Approximate control:

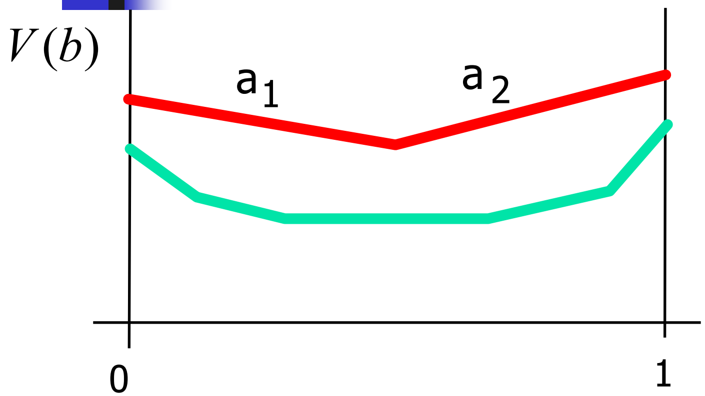
$$\hat{\pi}(b) = \arg\max_{a \in A} \left[R(b,a) + \gamma \sum_{o \in \Theta} P(o \mid b,a) \hat{V}(\tau(b,o,a)) \right]$$

- **Structured VF approximation methods**

$$V^*(s, b(. \mid o)) \approx \hat{V}(s, b(. \mid o)) \quad \forall s$$

Structured versions of methods in Hauskrecht (AAAI-97)

Fast informed bound method



Main properties:

- uses |A| linear functions (one linear function per action)
- upper-bounds the optimal function
- is computable efficiently (LP)

exact update ($|A||\Gamma_i|^{|\Theta|}$ linear functions)

$$V_{i+1}(b) = \max_{a \in A}[\sum_{s' \in S} b(s')R(s',a) + \gamma \sum_{o \in \Theta} \max_{\alpha_i \in \Gamma_i} \sum_{s' \in S} \sum_{s \in S} P(s,o|s',a)b(s')\alpha_i(s)]$$

$\leq$

new update ($|A|$ linear functions)

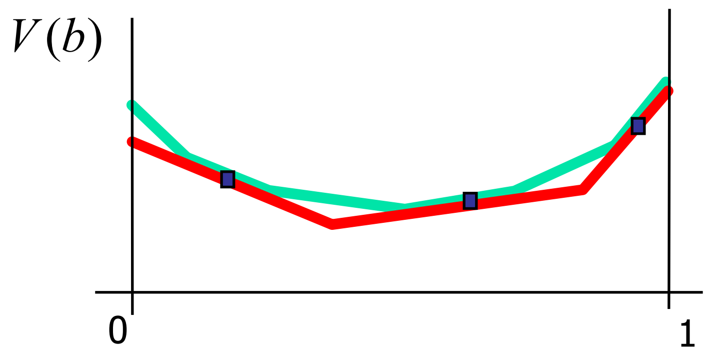
$$V_{i+1}(b) = \max_{a \in A}[\sum_{s' \in S} b(s')\underbrace{[R(s',a) + \gamma \sum_{o \in \Theta} \max_{\alpha_i \in \Gamma_i} \sum_{s \in S} P(s,o|s',a)\alpha_i(s)]}_{\alpha_{i+1}^{a}(s')}]$$

Grid (point) based linear function method

Grid-based approximations (Lovejoy 91)

- update value function values at |G| grid points
- **update derivatives (linear functions) at |G| grid points**

given a piecewise linear and convex V_i , we can compute efficiently the best linear function from V_{i+1} for any belief point



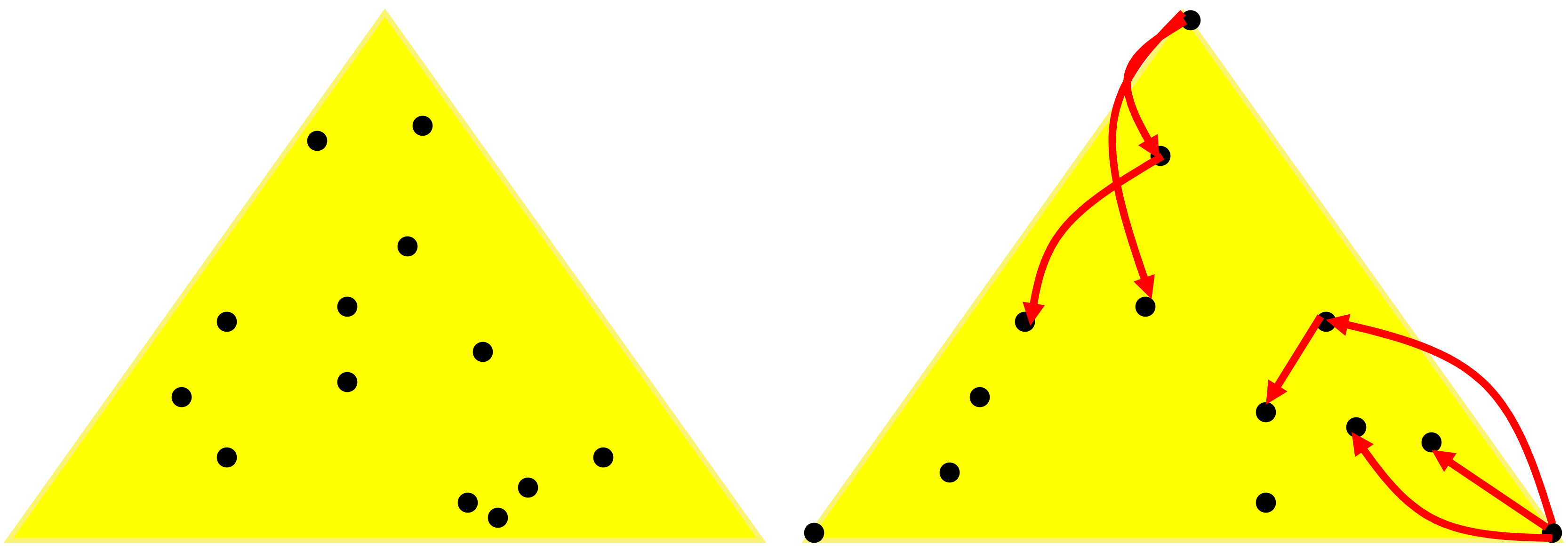
Properties

- efficient, approximation with $|G|$ linear functions
- lower-bounds the optimal solution

Grid (point) based approximation

Caveat: How to identify grid points?

We want value function approximations for all initial belief points



Random grids

Simulated grids from corners

Testing of the model

Testing and evaluation

Test the model on a "small" set of patient cases (12) with follow-ups (designed by a cardiologist)

Objectives

test model correctness, detect deficiencies, identify further refinements needed

Results of initial evaluation

- recommendations based on POMDP model mostly in agreement with a cardiologist (~85 %)
- disagreements caused by oversimplifications of the model (state description)

Patient case

	Current patient status	Recommended action
T0:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, negative Hx of PTCA	stress test no action medication
T1:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result positive, negative Hx of CABG, negative Hx of PTCA	PTCA stress test no action
T2:	no chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result normal, stress test result not available, negative Hx of CABG, Hx of PTCA	no action medication stress test
T3:	mild-moderate chest pain, positive resting EKG, normal ventricular function, acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	medication no action PTCA
T4:	mild-moderate chest pain, negative resting EKG, decreased ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG,Hx of PTCA	PTCA medication no action

Patient case

	Current patient status	Recommended action
T0:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, negative Hx of PTCA	stress test no action medication
T1:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result positive, negative Hx of CABG, negative Hx of PTCA	PTCA stress test no action
T2:	no chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result normal, stress test result not available, negative Hx of CABG, Hx of PTCA	no action medication stress test
T3:	mild-moderate chest pain, positive resting EKG, normal ventricular function, acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	medication no action PTCA
T4:	mild-moderate chest pain, negative resting EKG, decreased ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	PTCA medication no action

Patient case

	Current patient status	Recommended action
T0:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, negative Hx of PTCA	stress test no action medication
T1:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result positive, negative Hx of CABG, negative Hx of PTCA	PTCA stress test no action
T2:	no chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result normal, stress test result not available, negative Hx of CABG, Hx of PTCA	no action medication stress test
T3:	mild-moderate chest pain, positive resting EKG, normal ventricular function, acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	medication no action PTCA
T4:	mild-moderate chest pain, negative resting EKG, decreased ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	PTCA medication no action

Patient case

	Current patient status	Recommended action
T0:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, negative Hx of PTCA	stress test no action medication
T1:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result positive, negative Hx of CABG, negative Hx of PTCA	PTCA stress test no action
T2:	no chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result normal, stress test result not available, negative Hx of CABG, Hx of PTCA	no action medication stress test
T3:	mild-moderate chest pain, positive resting EKG, normal ventricular function, acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	medication no action PTCA
T4:	mild-moderate chest pain, negative resting EKG, decreased ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	PTCA medication no action

Patient case

	Current patient status	Recommended action
T0:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, negative Hx of PTCA	stress test no action medication
T1:	mild-moderate chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result not available, stress test result positive, negative Hx of CABG, negative Hx of PTCA	PTCA stress test no action
T2:	no chest pain, negative resting EKG, normal ventricular function, no acute MI, catheter result normal, stress test result not available, negative Hx of CABG, Hx of PTCA	no action medication stress test
T3:	mild-moderate chest pain, positive resting EKG, normal ventricular function, acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	medication no action PTCA
T4:	mild-moderate chest pain, negative resting EKG, decreased ventricular function, no acute MI, catheter result not available, stress test result not available, negative Hx of CABG, Hx of PTCA	PTCA medication no action

What did work

- **The model**
  - After defining many model parameters we got very reasonable behavior in many simulated patient case scenarios
- **The algorithms:**
  - Structured value function approximation methods (FIB and our point-based method) worked well for the IHD problem
  - Add-on: Smart cashing
  - A relatively small belief space for two hidden variables (9 configurations of hidden values)

What did not work well

- **Model acquisition:**
  - Too many parameters to define
  - Can Reinforcement Learning help us?

    Caveats: We do not have access to unlimited number of cases. We cannot choose the next action.
- **Repeated investigative actions:**
  - **Example:**
    - Stress test result was not obtained
    - A repeated stress test for a patient had the same outcome
  - Do not calculate the expected outcomes of actions for an individual from the population

Acknowledgments

- Hamish Fraser
- Peter Szolovits
- William Long