[PO]MDP in practice

The tiger problem

Outcomes

State

Action

	-100	**+ 10**
	+ 10	**-100**
	-1	**-1**

Observations

	0.50	0.50
	0.50	0.50
	0.85	0.15

	0.50	0.50
	0.50	0.50
	0.15	**0.85**

Transitions

0.50
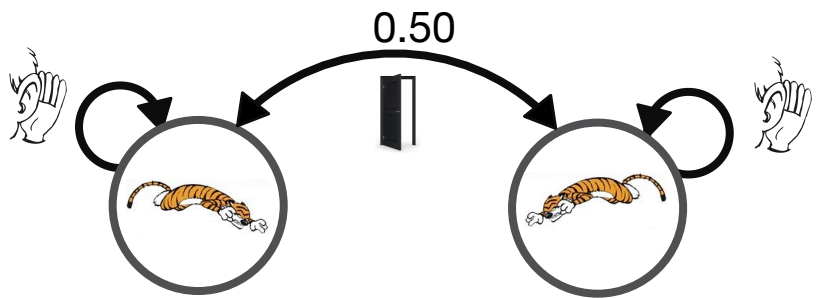
POMDP toolbox

- Download from
  - Official www.pomp.org http://cs.brown.edu/research/ai/pomdp/
  - Github https://github.com/lionel-rigoux/CPC-pomdp

- Compile the C code

- Edit the POMDP description file

- Run the executable or the the helper script

POMDP solution

alpha vectors (.alpha file)

// action
// v_1 v_2 v_3 ...
//
// ...

$b = [p(s=s_1)\ p(s=s_2)\ p(s=s_3)\ ...]$, sum to 1

$V(action, b) = v1\ p(s=s_1) + v2\ p(s=s_2) + v3\ p(s=s_3)$

alpha-vector

belief

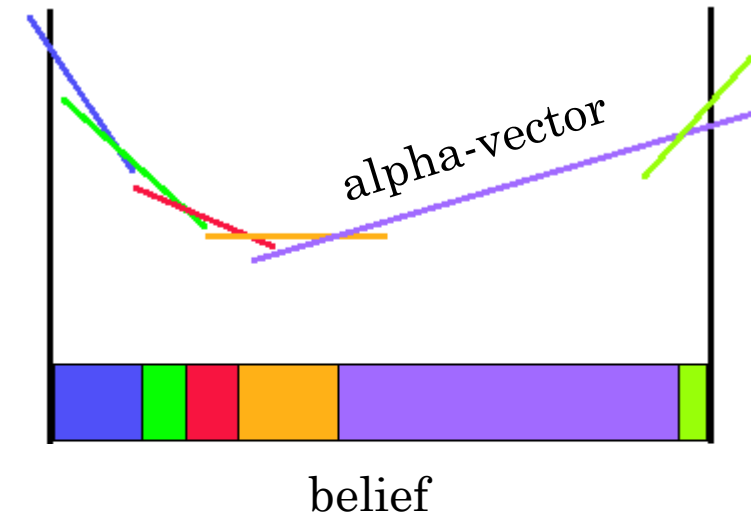Some actions can have multiple alpha vectors, some none!

POMDP simulation

internalized world
T(s'|s,a), O(o|s,a), R(a,s)

real world
T(s'|s,a), O(o|s,a), R(a,s)

INITIALIZE

a. find optimal policy
b. pick initial belief

a. pick initial state

ITERATE

1. chose best action given current belief   a

2. change state depending on action

3. get reward depending on state and action

o   4. get observation depending on state and action

5. update belief following given action, observation and internal model

POMDP summary

The toolbox allows to:

- Compute the optimal policy for a given representation of the world

- Compute the utility/value associated with beliefs and actions

- Simulate the effect of a false representation of the world on complex decision making and belief updating

It can not:

- capture how we learn the policy

- capture how representation of the world are learned

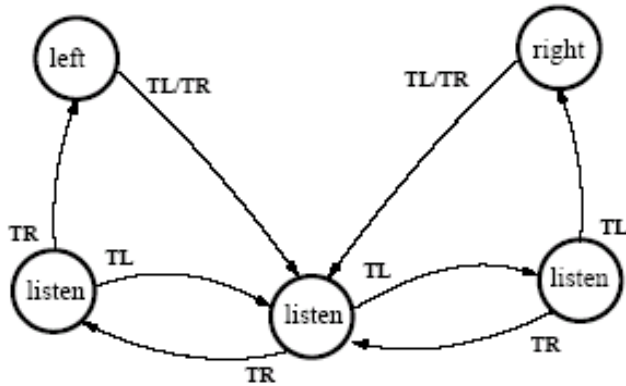- be used to fit experimental data (infer on hidden representations)
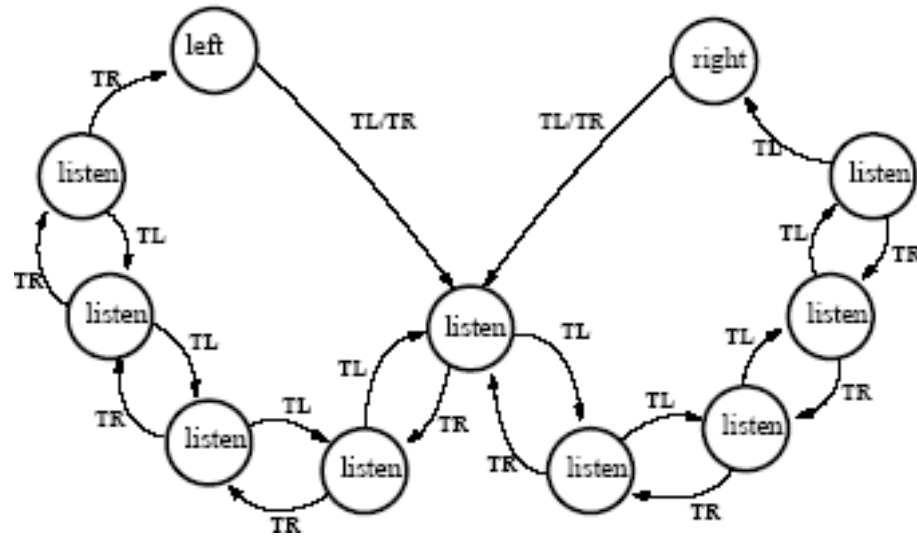
Questions?

POMDP solution

belief MDP (.pg file)
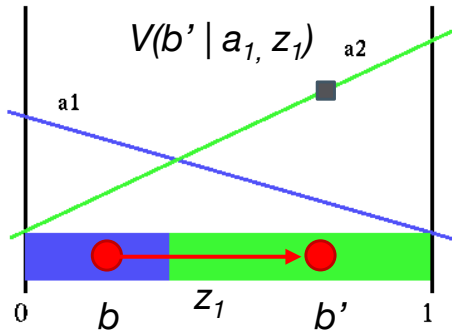

$p(o_{TL} \mid \text{listen, TL}) = 0.99$
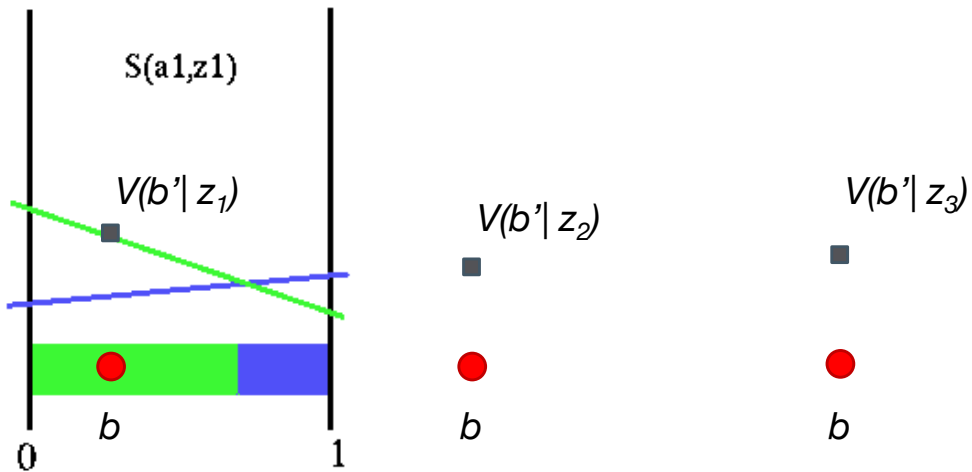
$p(o_{TL} \mid \text{listen, TL}) = 0.75$

Resolution

belief value $V(b) = r(b) + \gamma\, V(b')$

immediate reward

$V(b' \mid a_1,\, z_1)$ a2

a1

0   b z_1 b' 1

future value $V(b') = \sum V(b'|o)\, p(o|b)$

S(a1,z1)

$V(b'|\, z_1)$

$V(b'|\, z_2)$

$V(b'|\, z_3)$

b b b

0   1

z1

z2

z3

0   1