

[PO]MDP in practice

# The tiger problem



# The tiger problem



# The tiger problem



# The tiger problem



# The tiger problem



# The tiger problem



# The tiger problem





# The tiger problem



IF ANYTHING BAD HAPPENS,  
IT'S NOT MY FAULT. IT'S FATE.








# The tiger problem










# Outcomes

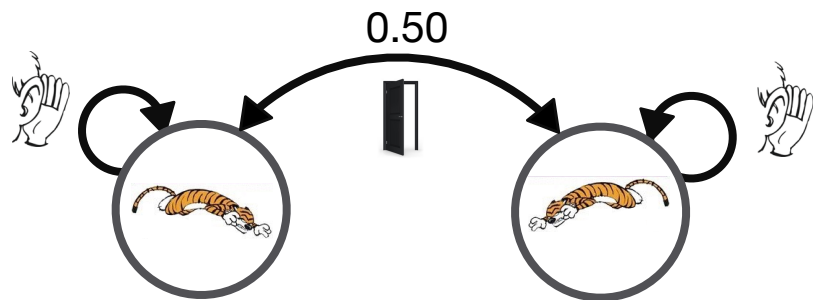
|        |   | State   |   |
|--------|---|---|---|
|        |   |  |  |
| Action |  | -100  | + 10  |
|        |  | + 10  | -100  |
|        |  | -1  | -1  |








# Outcomes

|        |   | State   |   |
|--------|---|---|---|
|        |   |  |  |
| Action |  | -100  | + 10  |
|        |  | + 10  | -100  |
|        |  | -1  | -1  |







# Transitions



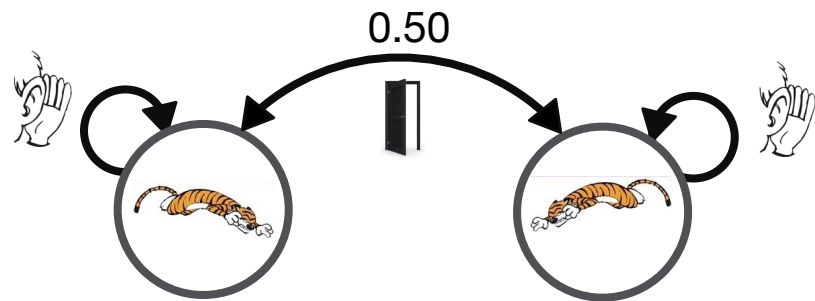
## Outcomes





|        |   | State   |   |
|--------|---|---|---|
|        |   |  |  |
| Action |  | -100  | + 10  |
|        |  | + 10  | -100  |
|        |  | -1  | -1  |

## Observations

|   |   |  |  |
|---|---|---|---|
|  |  | 0.50  | 0.50  |
|   |  | 0.50  | 0.50  |
|   |  | <b>0.85</b>   | 0.15  |

## Transitions



|   |   |      |             |
|---|---|------|-------------|
|  |    | 0.50 | 0.50        |
|   |   | 0.50 | 0.50        |
|   |  | 0.15 | <b>0.85</b> |



# POMDP toolbox

- Download from  
Official [www.pomp.org](http://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



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$

belief value

immediate reward

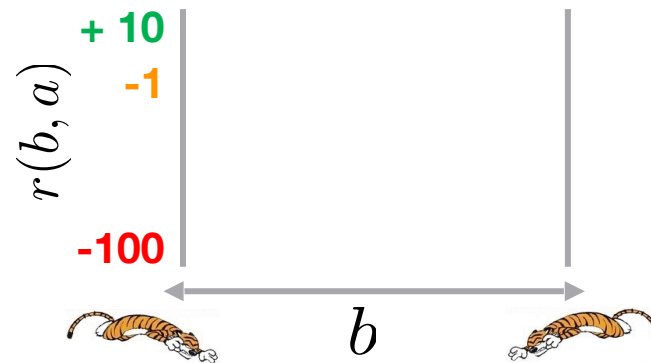
next belief value





# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

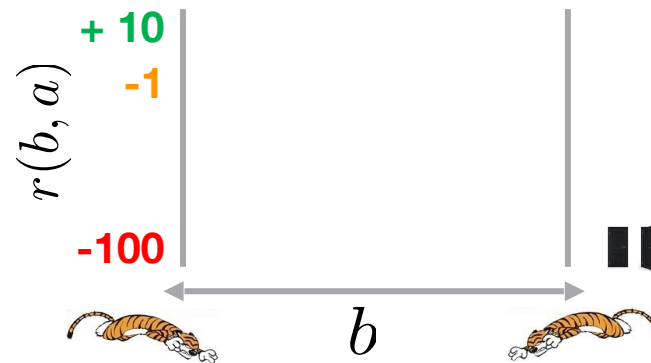
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

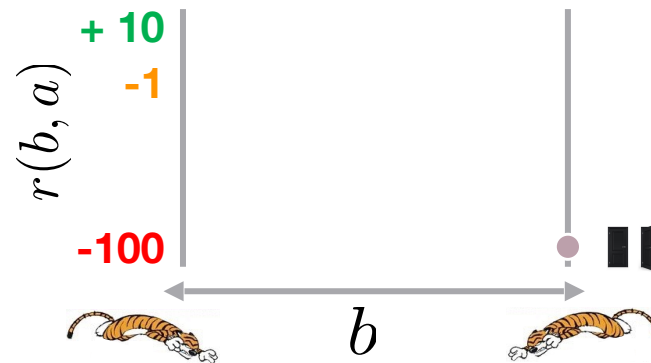
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

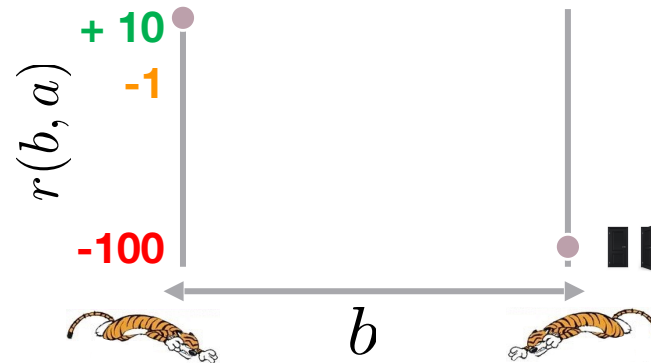
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

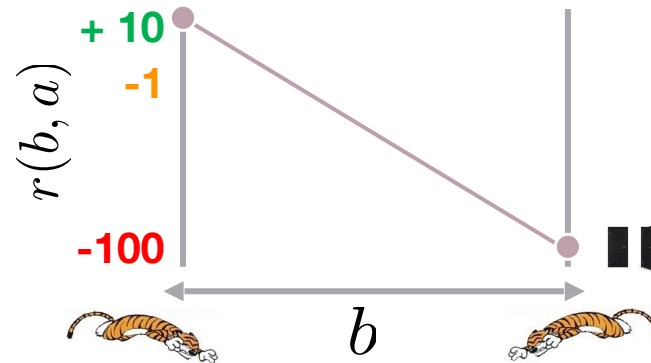
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

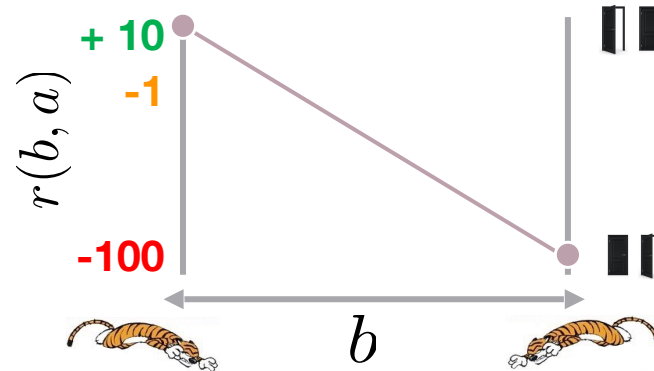
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

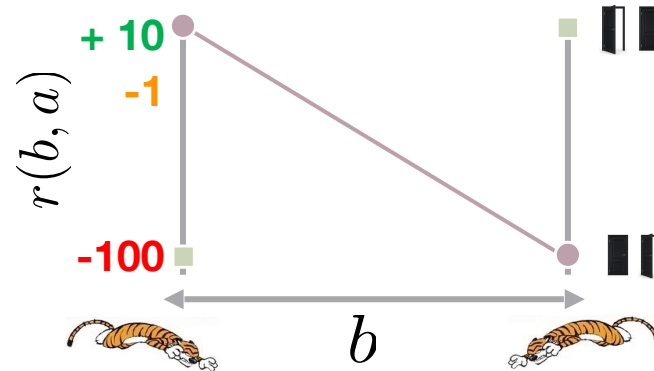
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

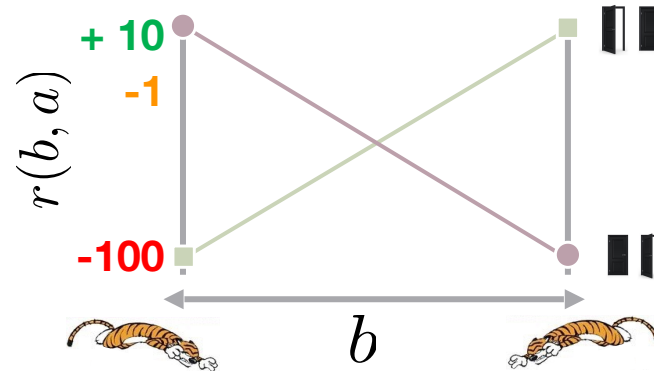
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

immediate reward

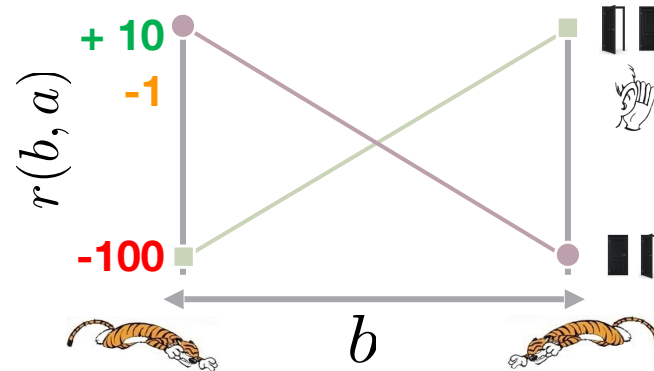
next belief value





# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

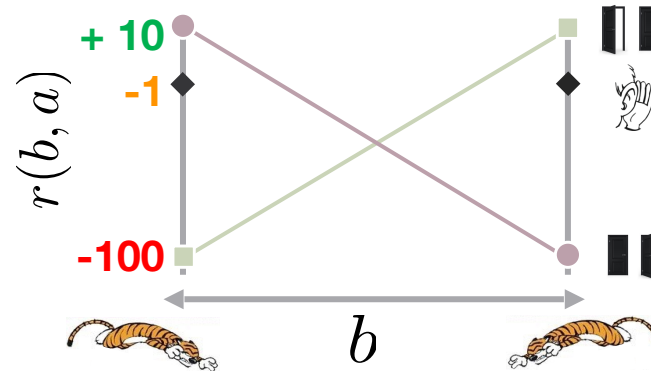
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

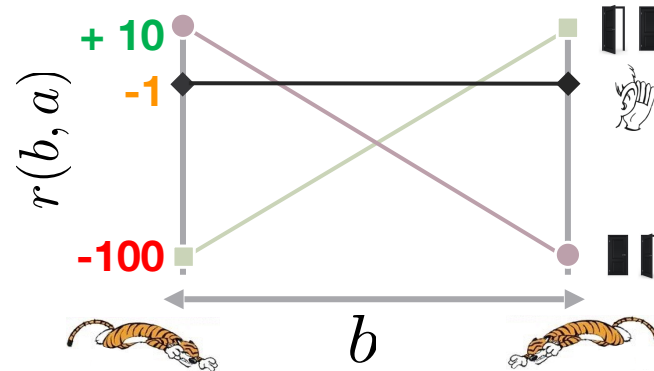
immediate reward

next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

immediate reward

next belief value



# Resolution

**current belief value**

**next belief value**



# Resolution

$$V(b, a) = r(b, a)$$

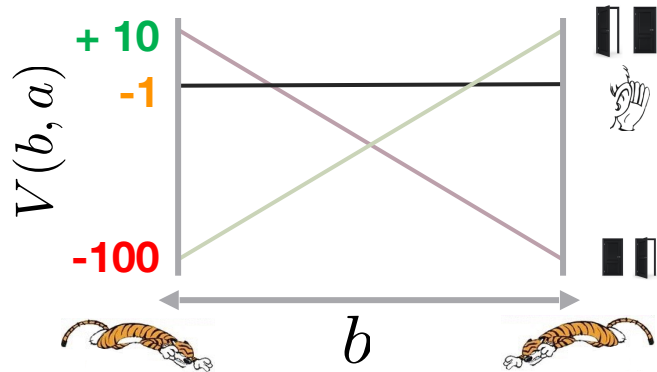
**current belief value**

**next belief value**



# Resolution

$$V(b, a) = r(b, a)$$



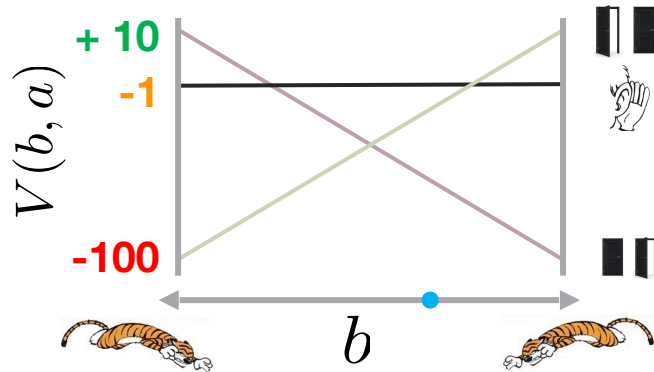
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



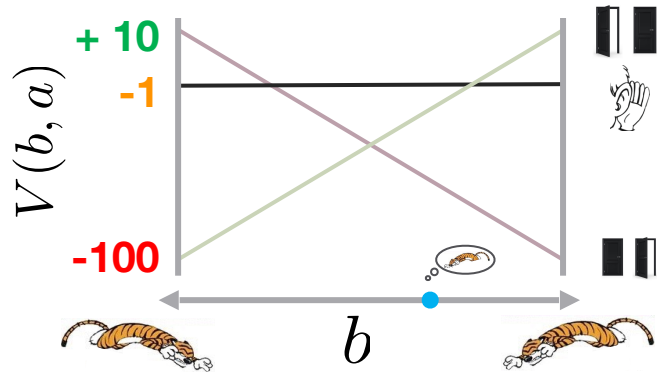
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



current belief value

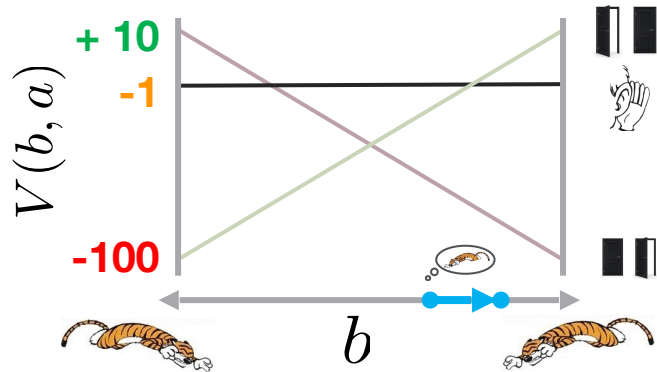
next belief value





# Resolution

$$V(b, a) = r(b, a)$$



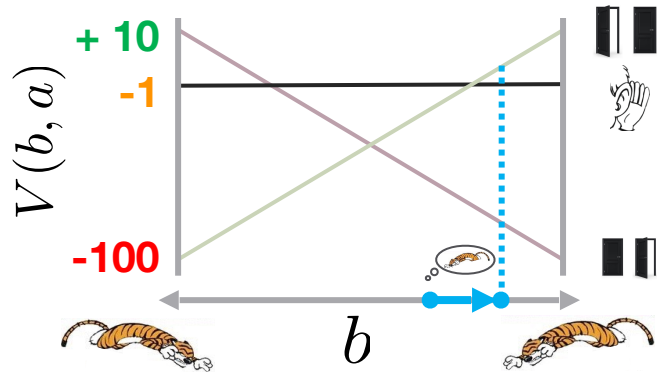
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



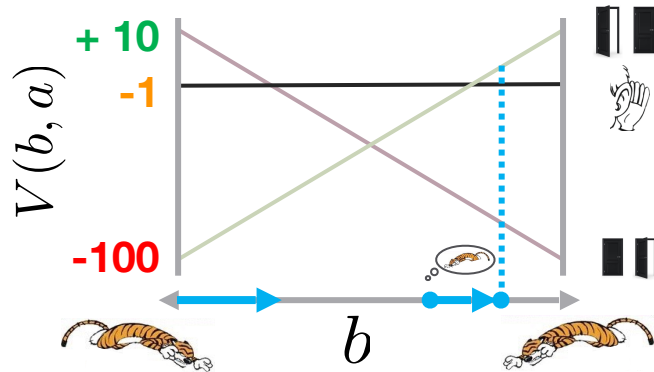
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



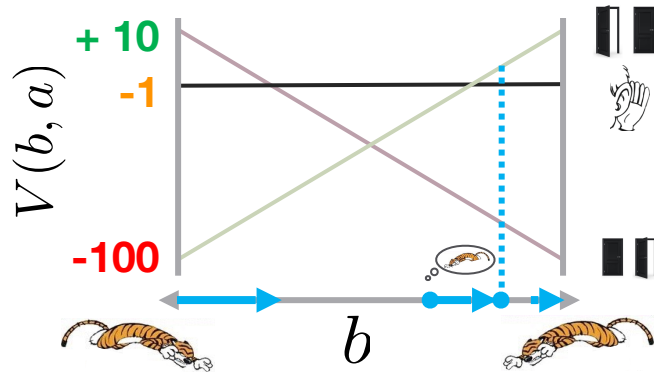
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



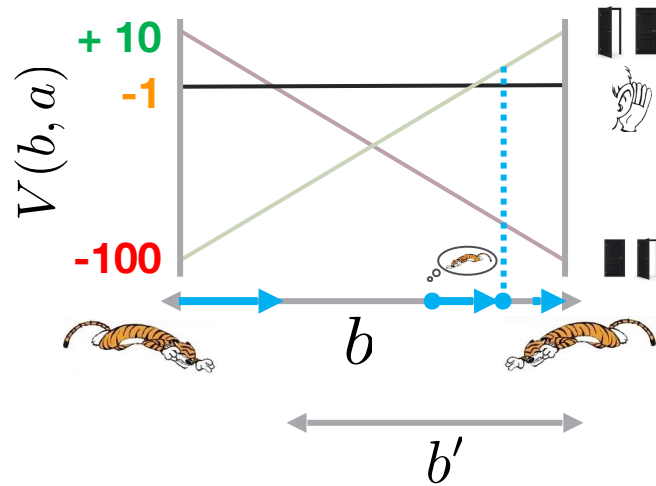
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



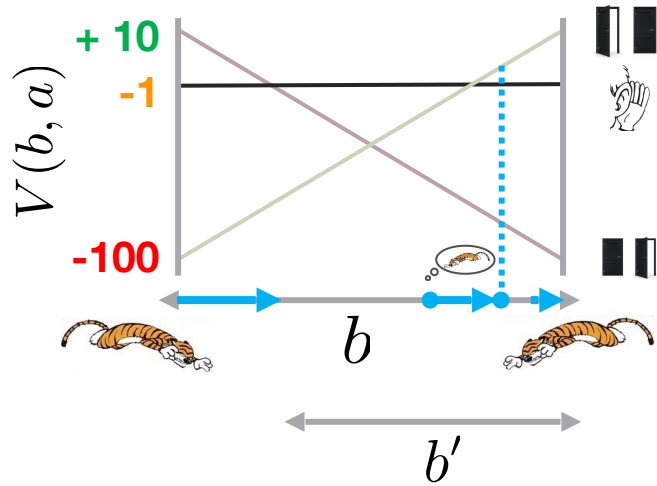
current belief value

next belief value



# Resolution

$$V(b, a) = r(b, a)$$



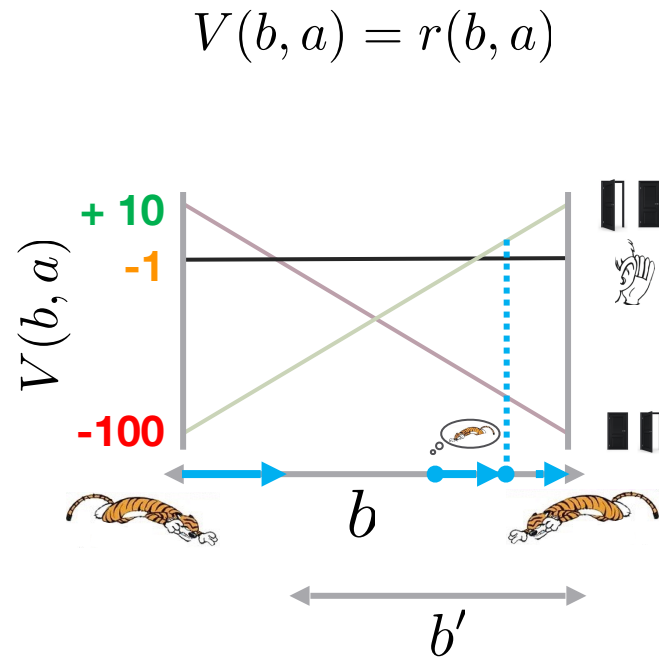
current belief value

$$V(b', a)$$

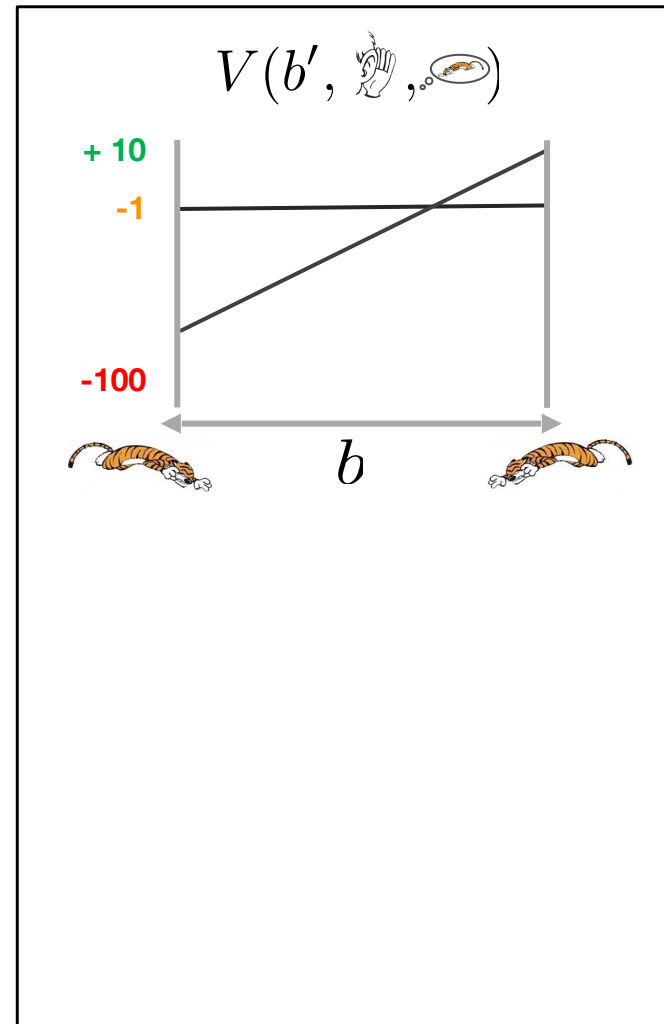
next belief value



# Resolution



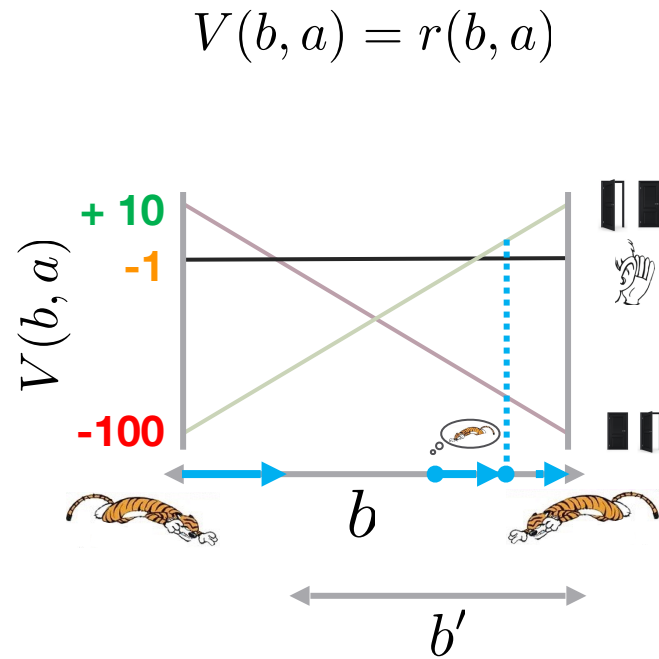
current belief value



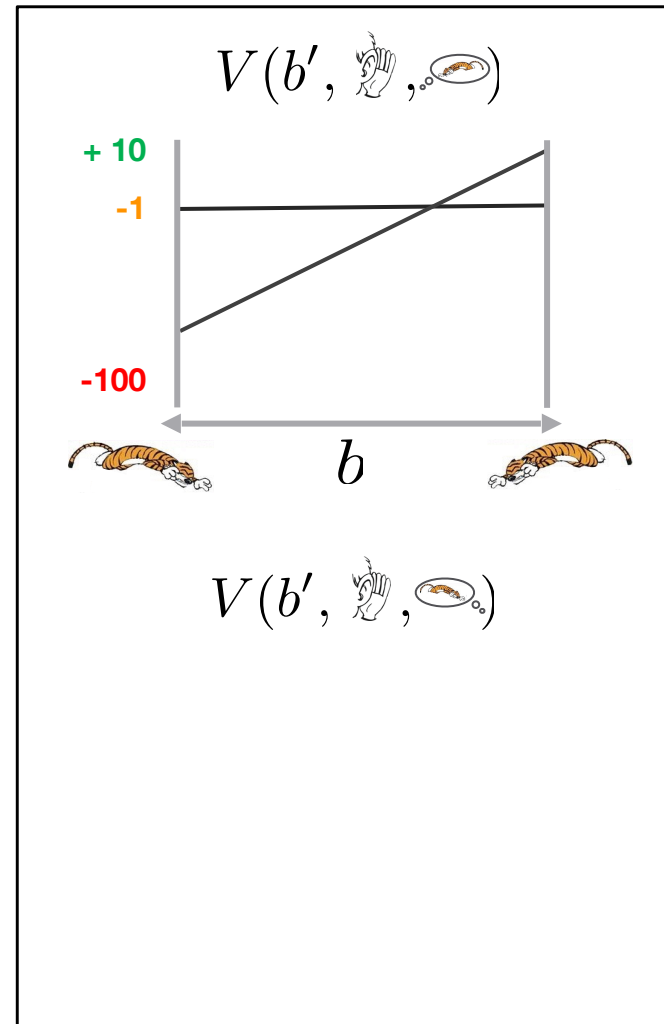
next belief value



# Resolution



current belief value

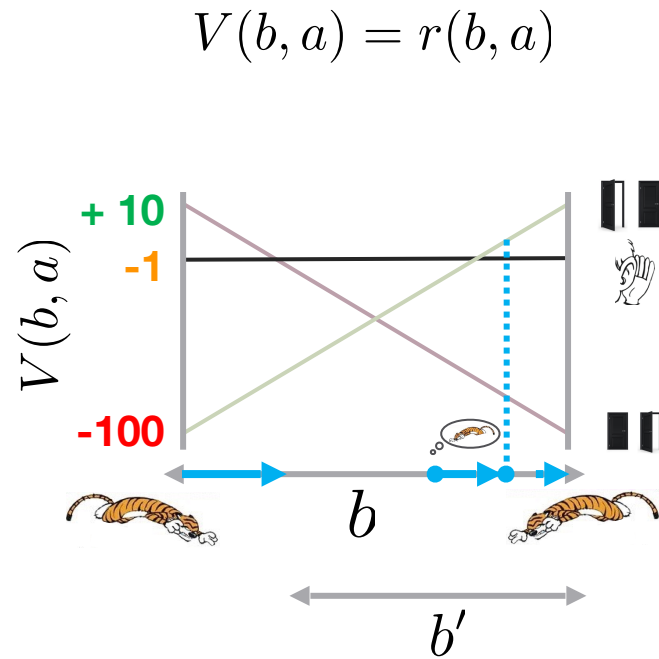


next belief value

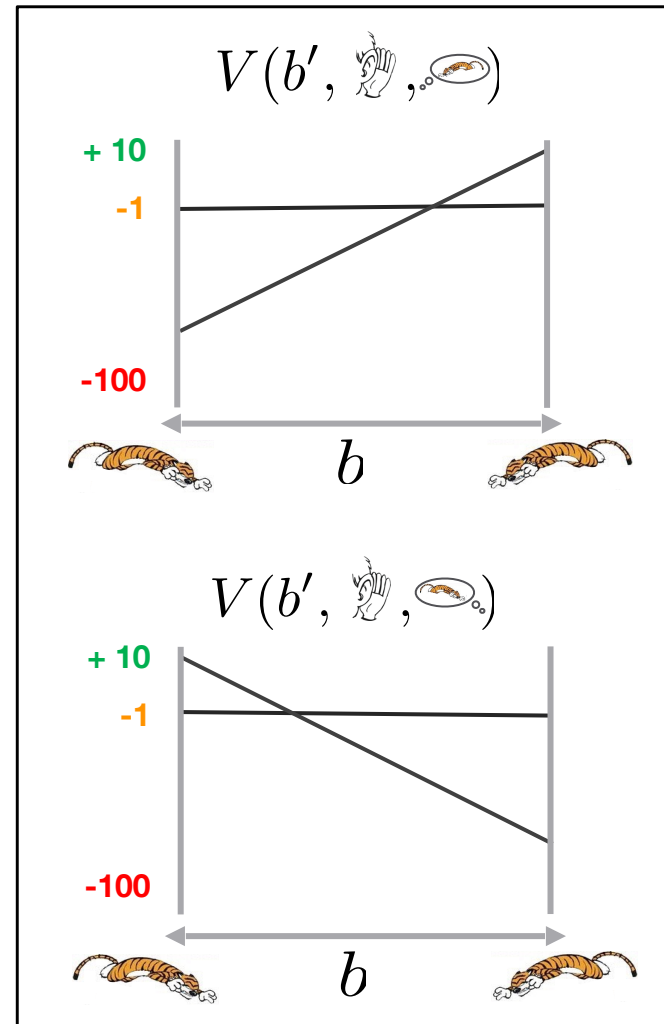




# Resolution



current belief value

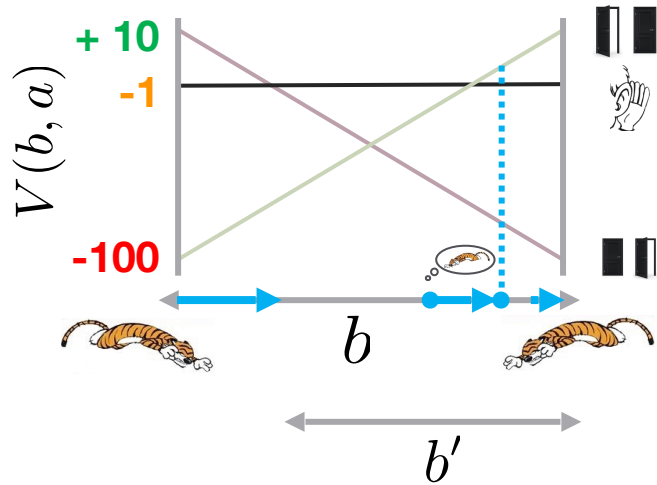


next belief value

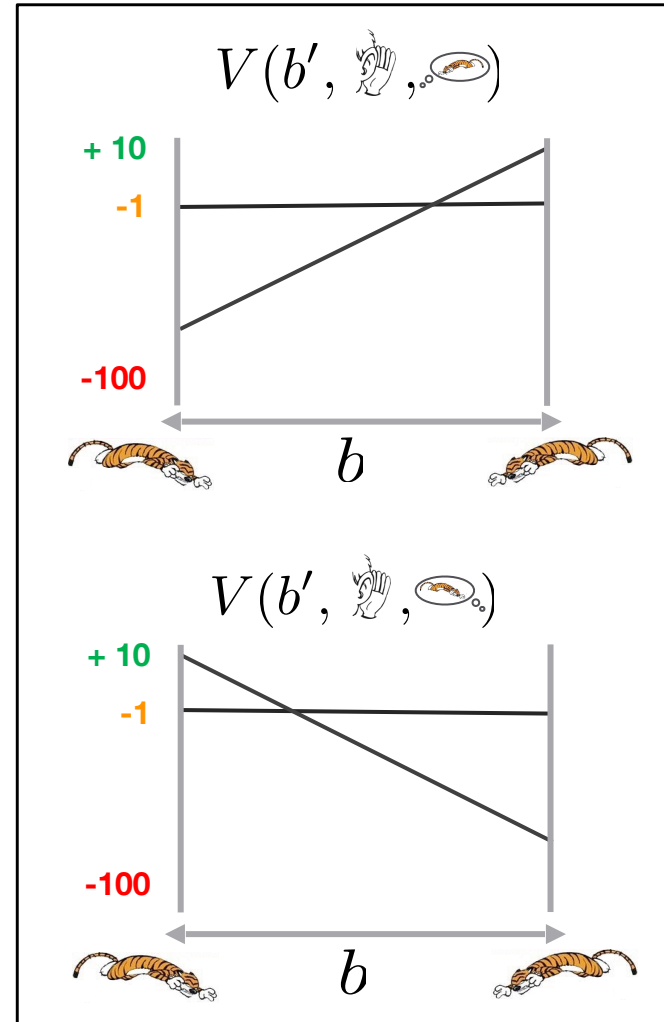


# Resolution

$$V(b, a) = r(b, a)$$



current belief value



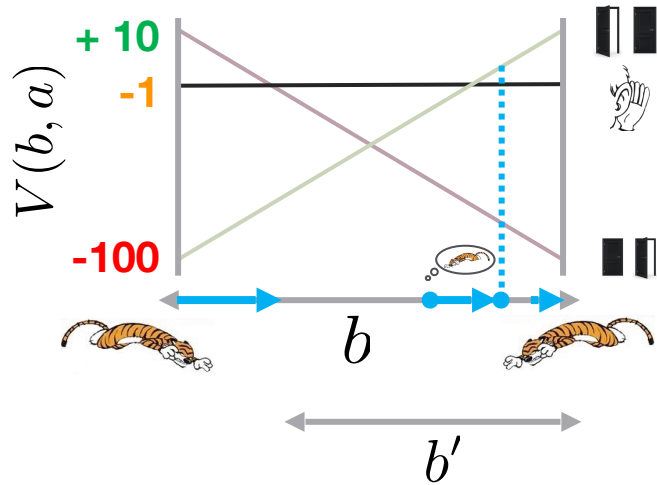
$$V(b') = \sum V(b'|o)p(o|b)$$

next belief value

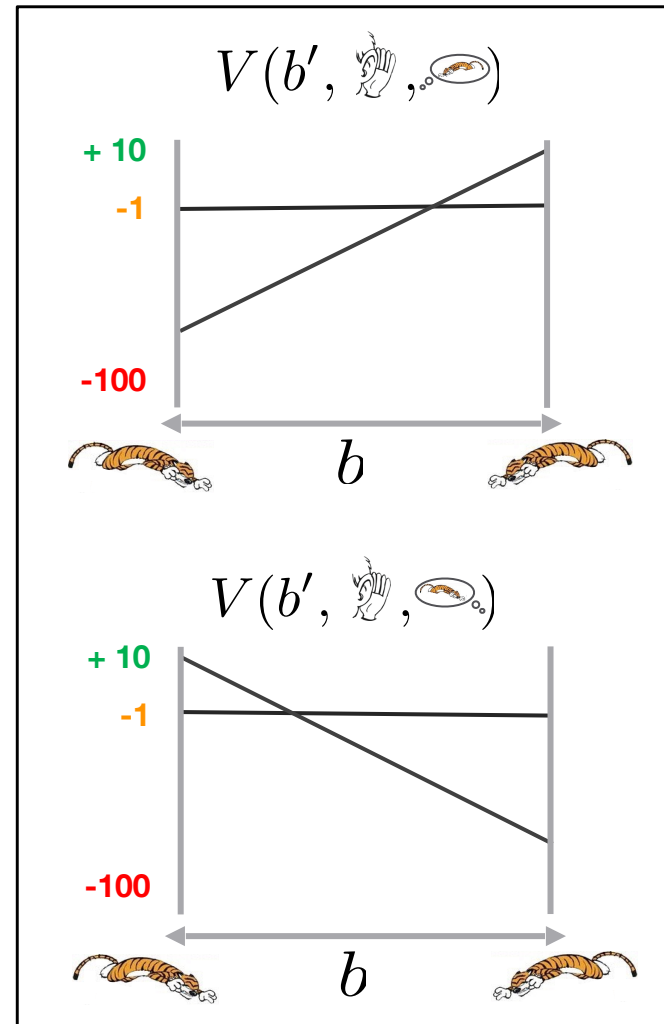


# Resolution

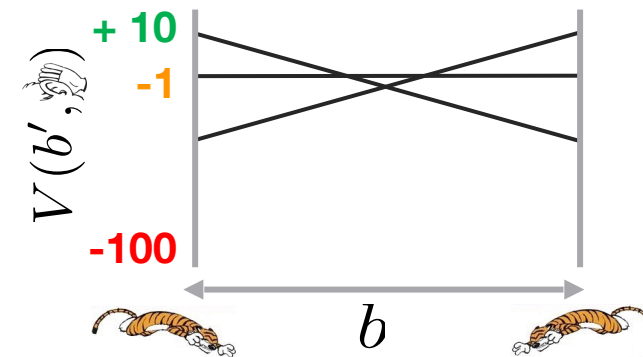
$$V(b, a) = r(b, a)$$



current belief value



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

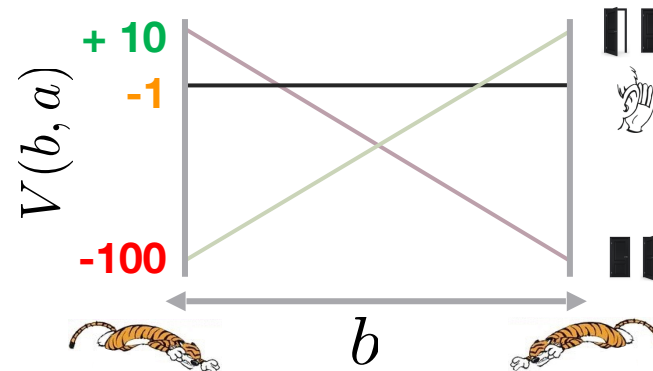


next belief value



# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

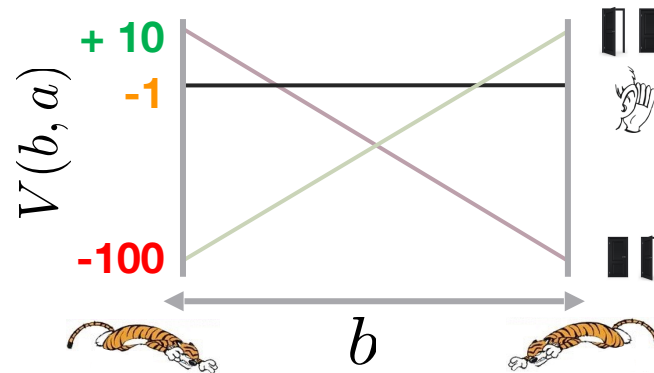
immediate reward

next belief value



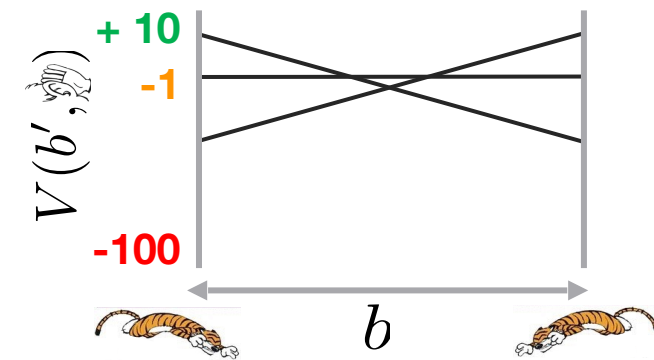
# Resolution

**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value

immediate reward

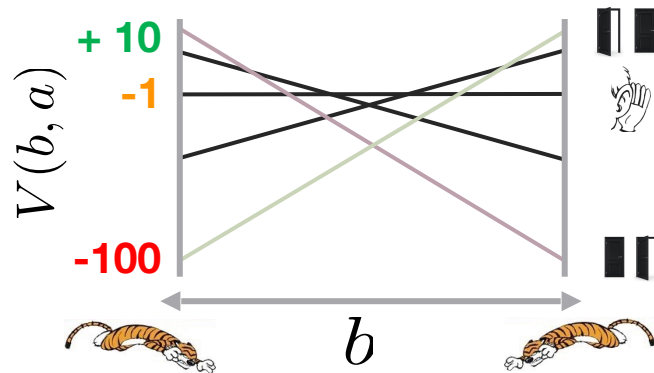


next belief value

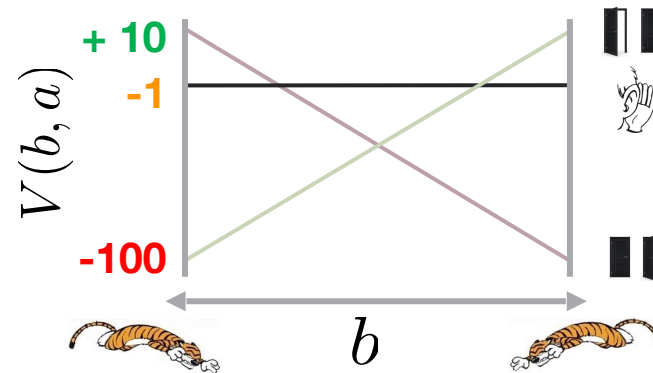


# Resolution

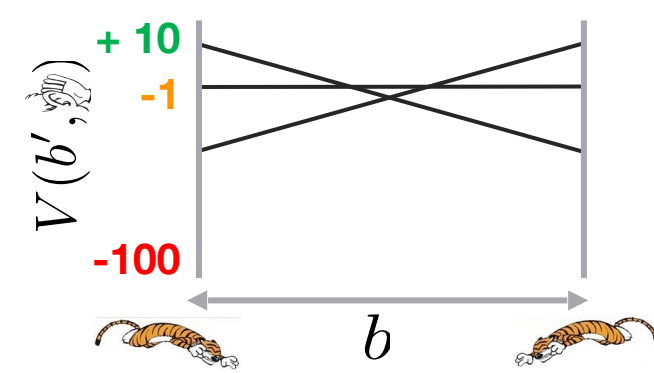
**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value



immediate reward

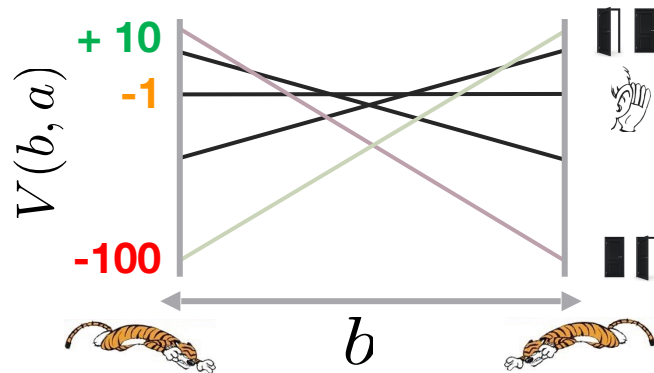


next belief value

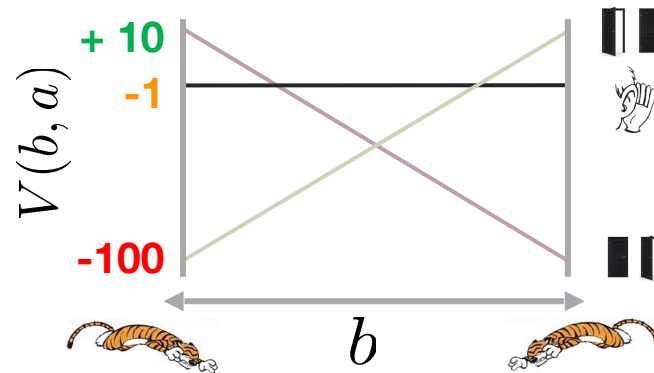


# Resolution

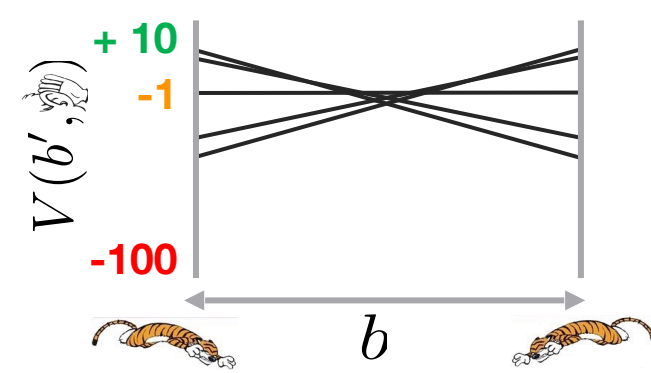
**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value



immediate reward

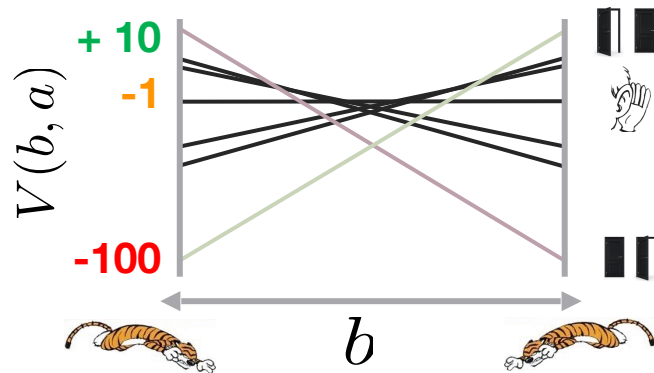


next belief value

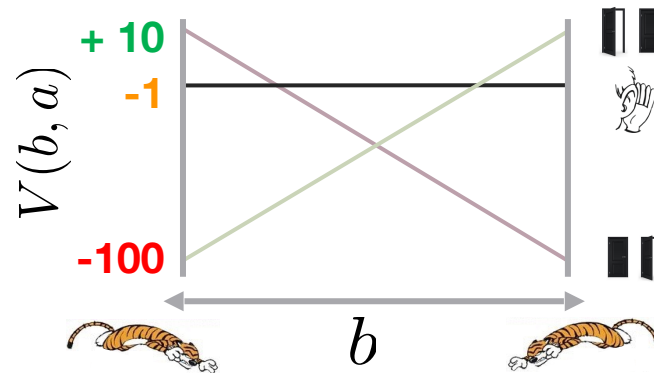


# Resolution

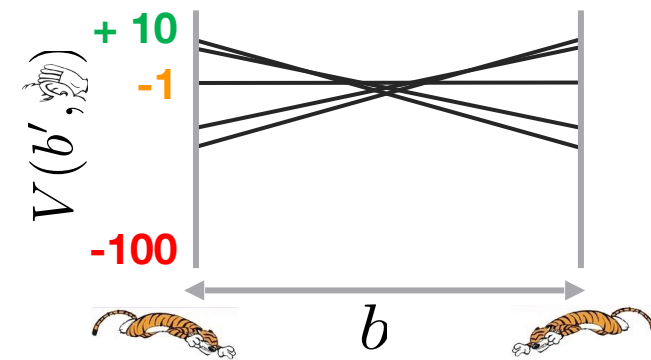
**Bellman principle**  $V(b, a) = r(b, a) + \gamma V(b', a)$



belief value



immediate reward



next belief value





# POMDP solution

**alpha vectors (.alpha file)**

```
// action  
// v1 v2 v3 ...  
//  
// ...
```

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

$V(\text{action}, b) = v_1 \ p(s=s_1) + v_2 \ p(s=s_2) + v_3 \ p(s=s_3)$

Some actions can have multiple alpha vectors, some none!



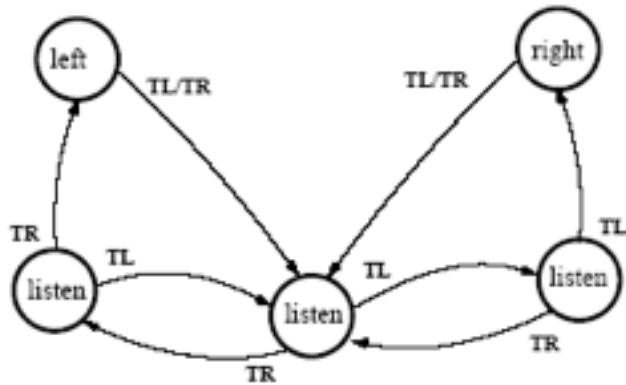
# POMDP solution

**belief MDP (.pg file)**



# POMDP solution

belief MDP (.pg file)

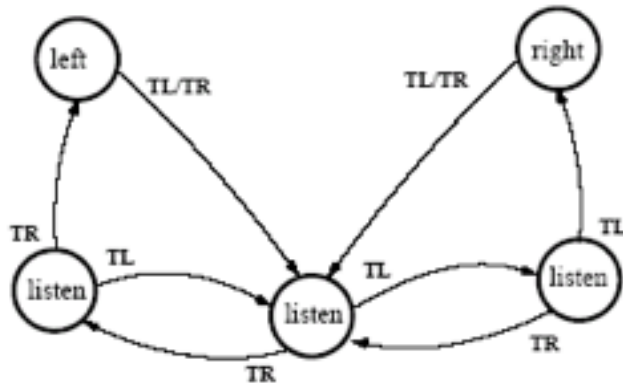


$$p(o_{TL} \mid \text{listen}, TL) = 0.99$$

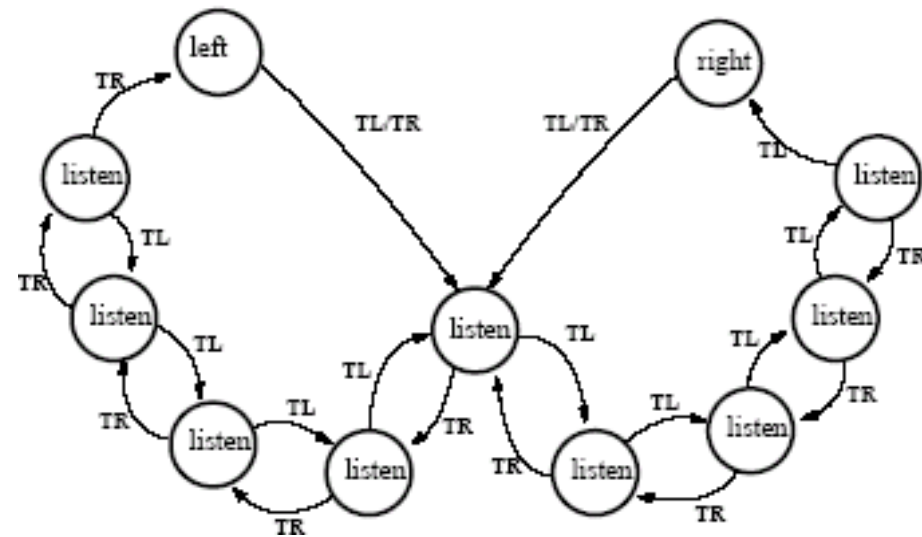


# POMDP solution

belief MDP (.pg file)

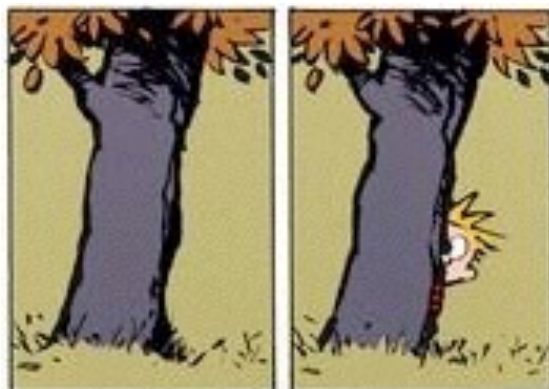


$$p(o_{TL} \mid \text{listen}, TL) = 0.99$$



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





# calvin and Hobbes

by BILL WATSON

drawn by BILL WATSON

10/20



DING DONG



C'MON, C'MON...  
ANSWER THE  
DOOR!



DING DONG  
DING DONG  
DING DONG  
DING DONG



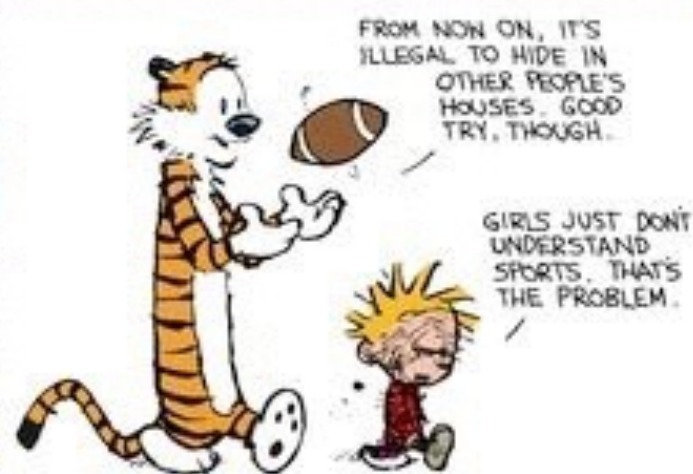
WHAT'S WRONG WITH YOU PEOPLE?  
DON'T YOU ANSWER YOUR DOOR??



OH, IT FIGURES IT'S YOU.  
MOM SAYS OUR DOORBELL  
IS NOT A TOY. GO AWAY.



HOPELESS



FROM NOW ON, IT'S  
ILLEGAL TO HIDE IN  
OTHER PEOPLE'S  
HOUSES. GOOD  
TRY, THOUGH.

GIRLS JUST DON'T  
UNDERSTAND  
SPORTS. THAT'S  
THE PROBLEM.

