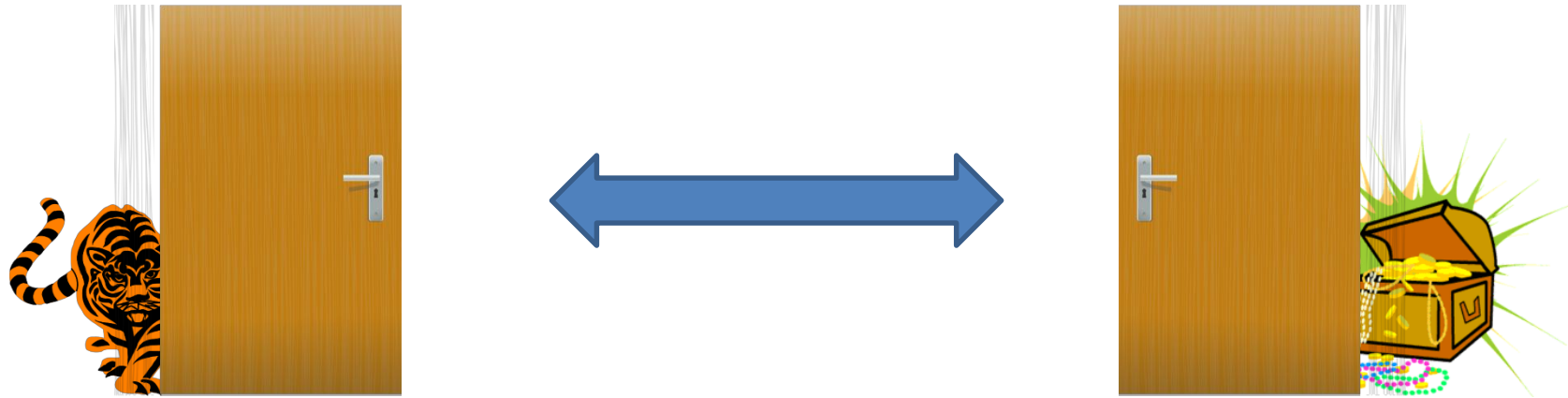
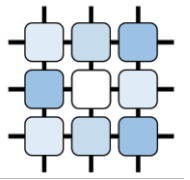


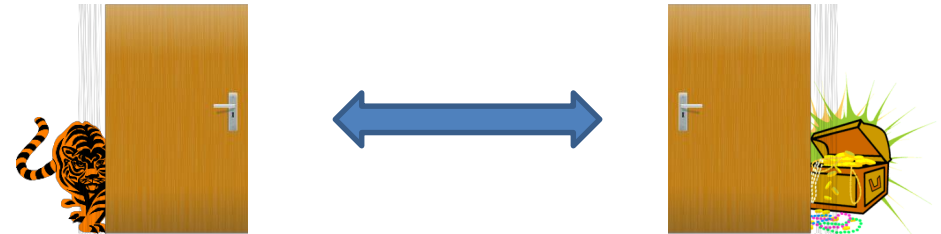
The Tiger Problem



- states = {tiger-left, tiger-right}
- actions = {listen, open-left, open-right}
 - transitions: no change (listen), restart (open-right, open-left)
- observations = {hear-tiger-left (TL), hear-tiger-right (TR)}
- rewards: surprised by tiger, found treasure, listening

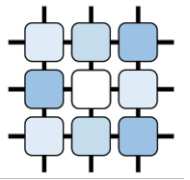


The Tiger Problem (transition prob.)

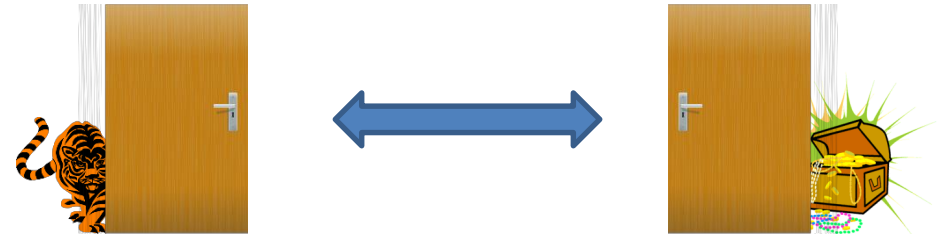


listen →	tiger-left	tiger-right
tiger-left	1.0	0.0
tiger-right	0.0	1.0

open-left/right →	tiger-left	tiger-right
tiger-left	0.5	0.5
tiger-right	0.5	0.5

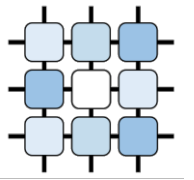


The Tiger Problem (observation prob.)

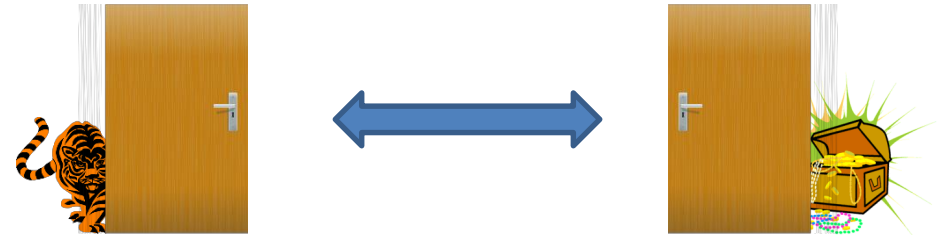


listen →	hear TL	hear TR
tiger-left	0.85	0.15
tiger-right	0.15	0.85

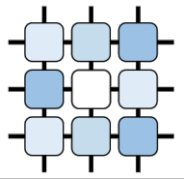
open-left/right →	hear TL	hear TR
tiger-left	0.5	0.5
tiger-right	0.5	0.5



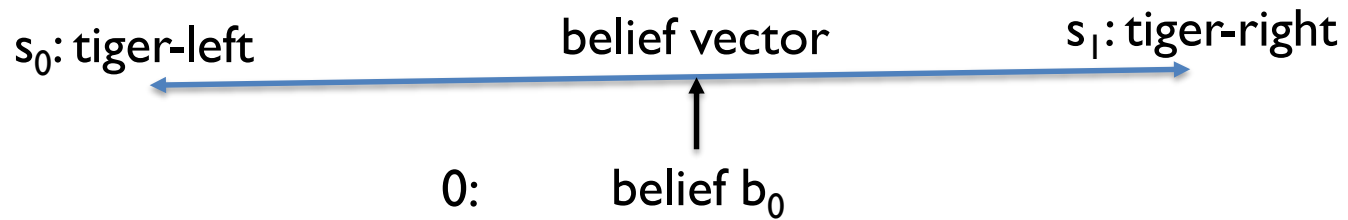
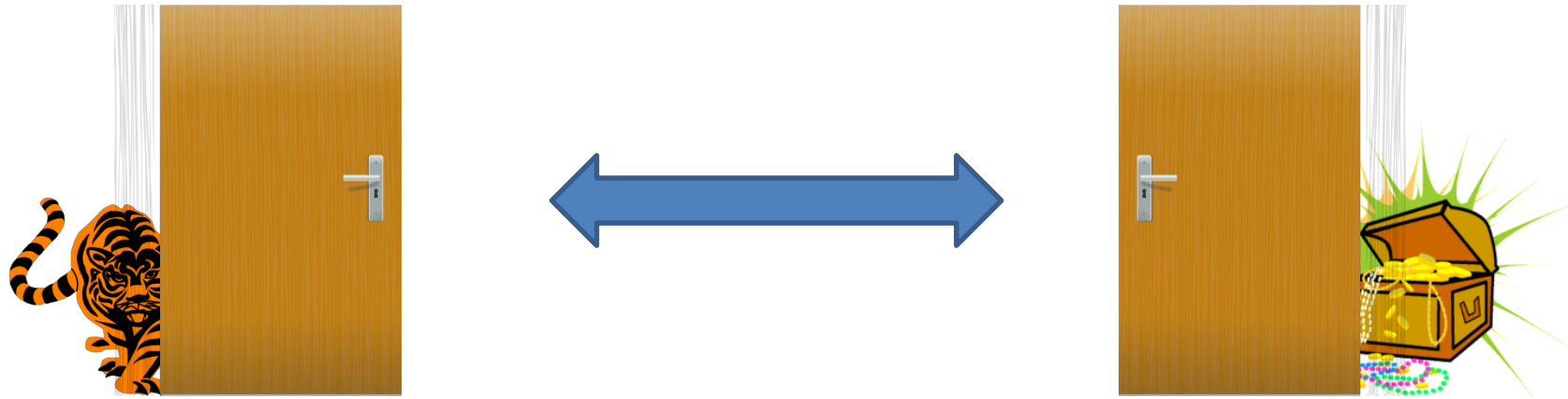
The Tiger Problem (immediate reward)

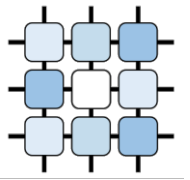


reward	tiger-left	tiger-right
listen	-1	-1
open-left	-100	+10
open-right	+10	-100

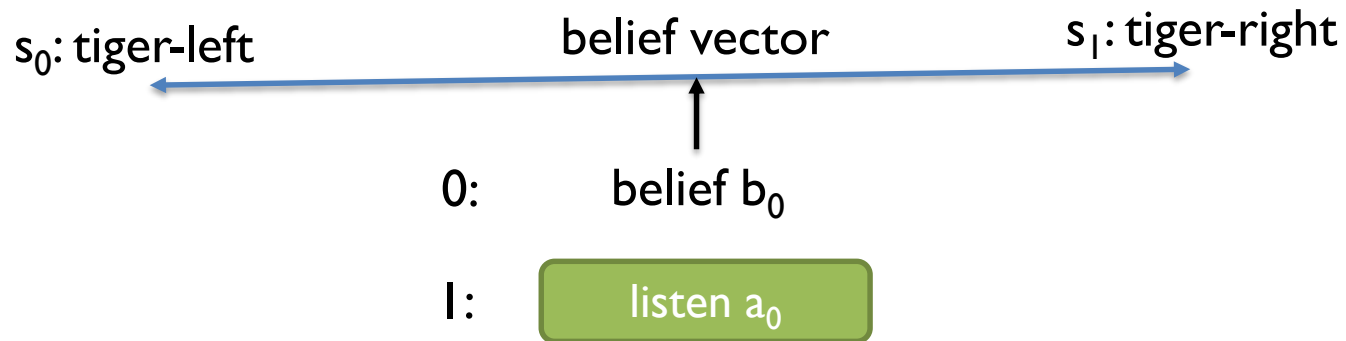
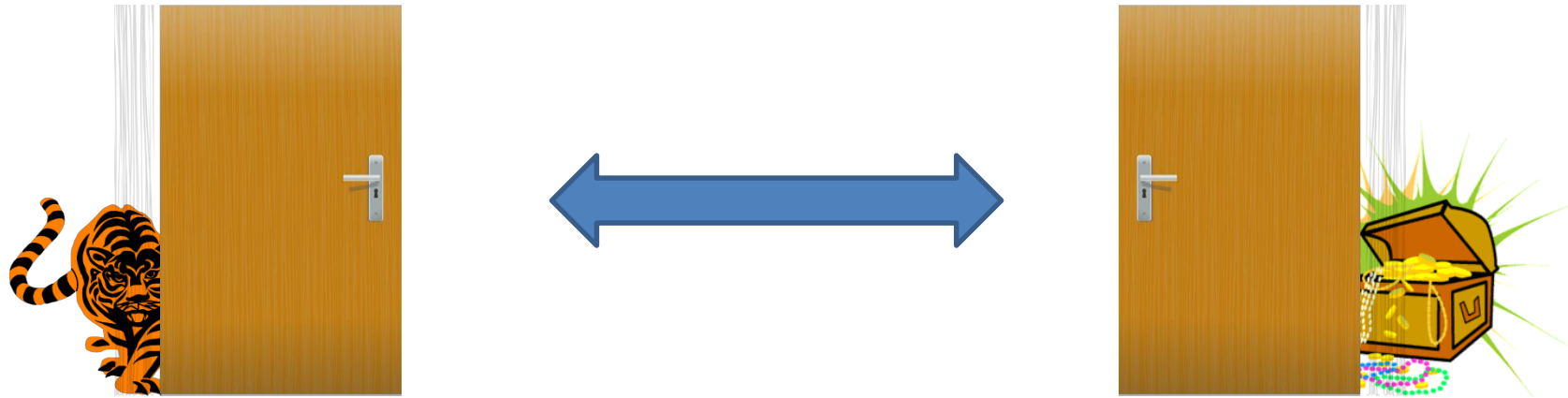


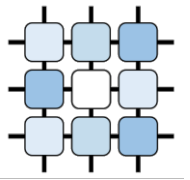
The Tiger Problem (belief update)



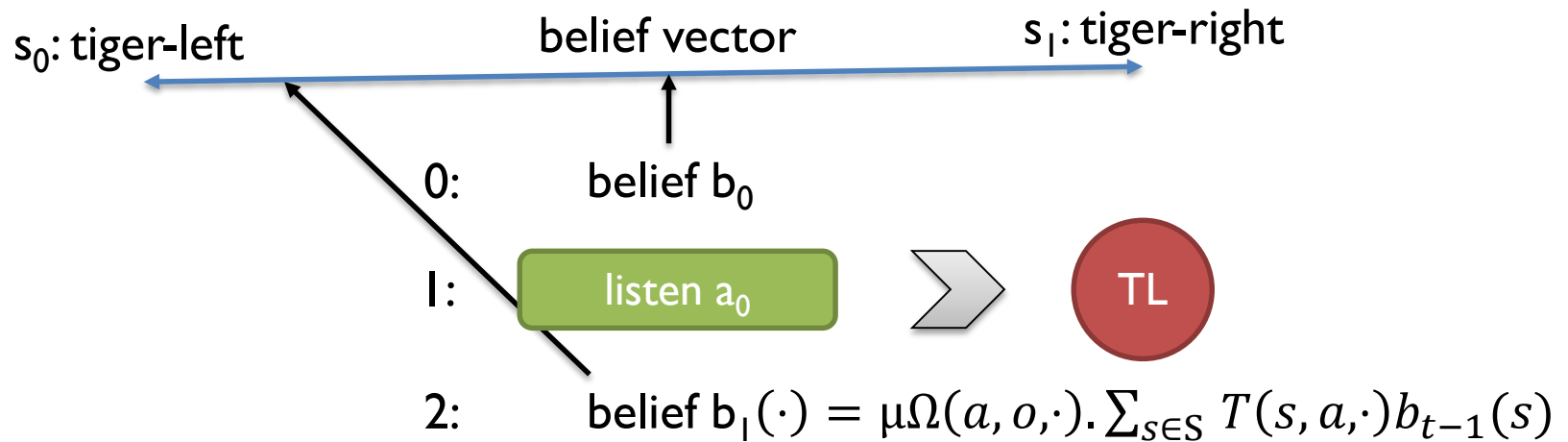
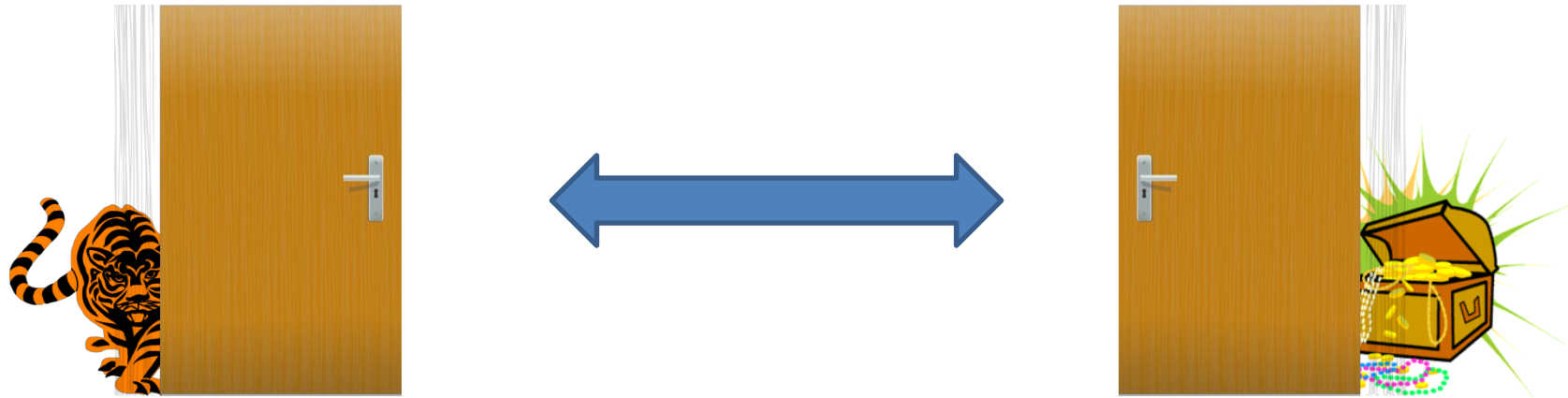


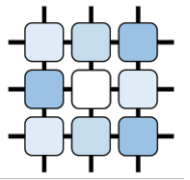
The Tiger Problem (belief update)



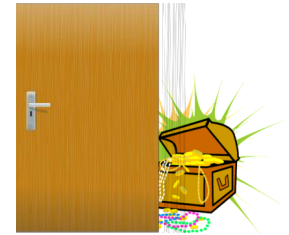
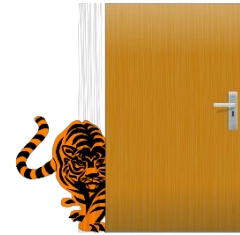


The Tiger Problem (belief update)





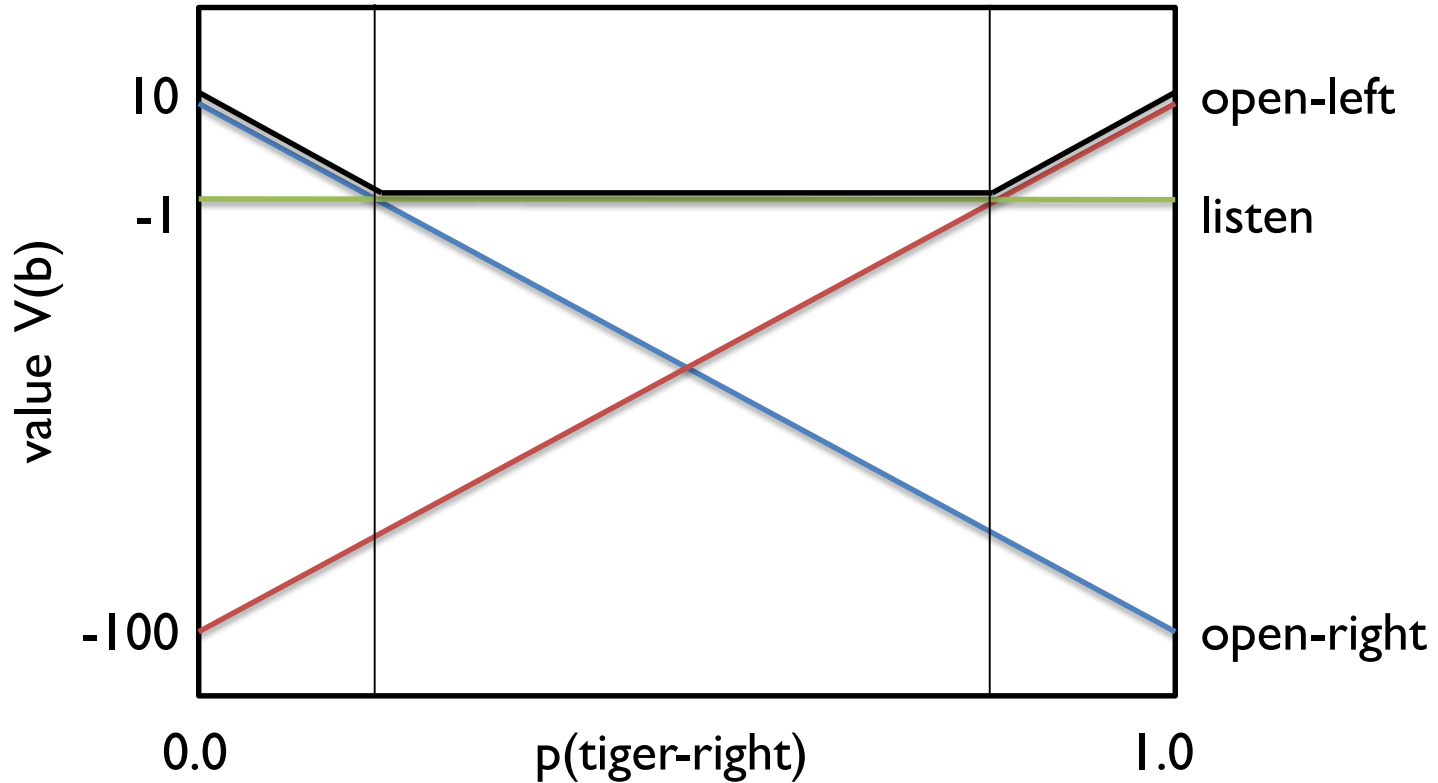
The Tiger Problem (1-step opt. policy)

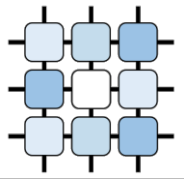


open-right
[0.0, 0.1]

listen
[0.1, 0.9]

open-left
[0.9, 1.0]





The Tiger Problem (2-step opt. policy)

