

Public Belief States

MIT 6.S890; October 31, 2024

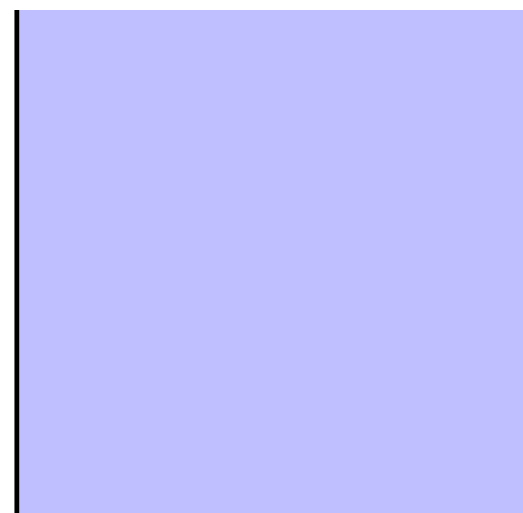
Why do we like symmetric-information games?

Why do we like symmetric-information games?

Backward induction!

Why do we like symmetric-information games?

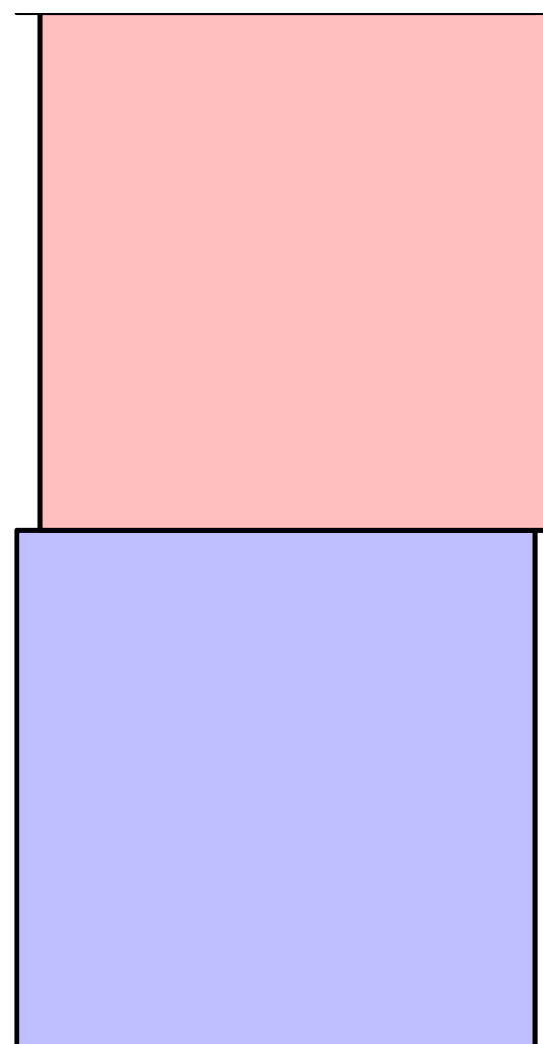
Backward induction!



s^{T-1}

Why do we like symmetric-information games?

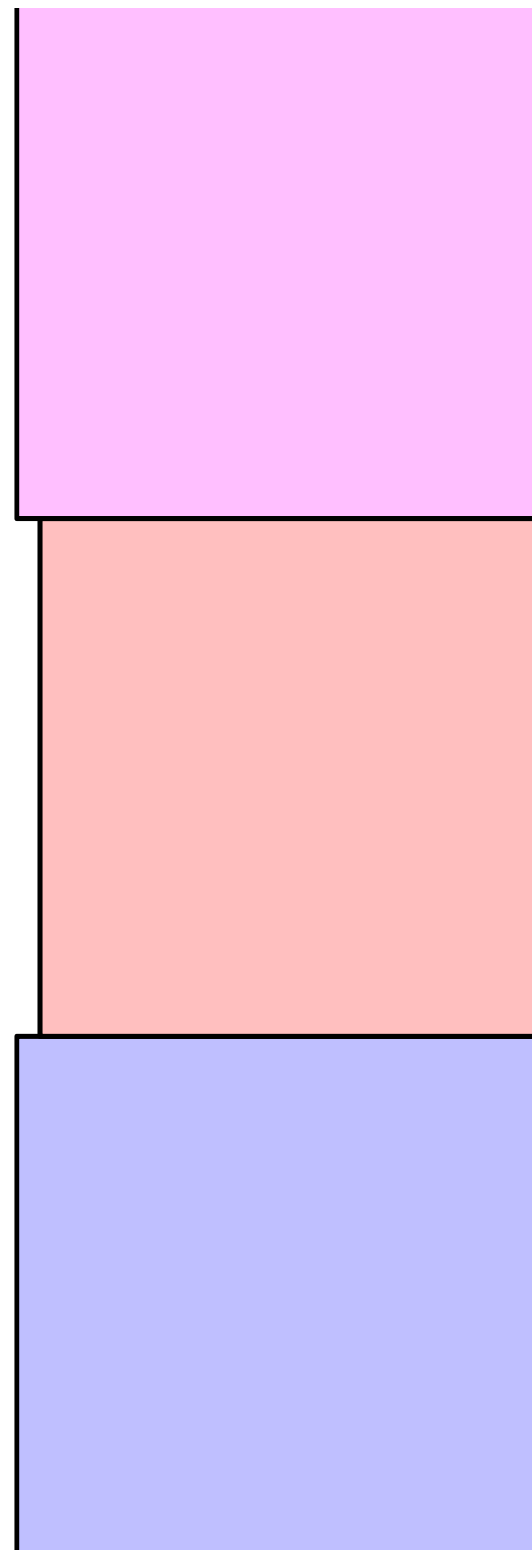
Backward induction!



s^{T-1}

Why do we like symmetric-information games?

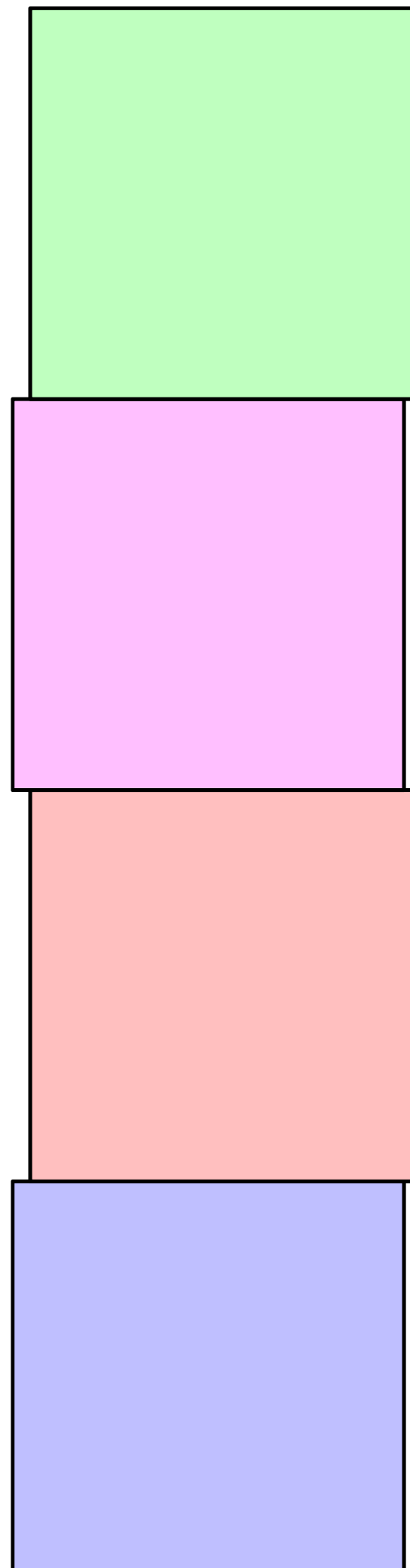
Backward induction!



s^{T-1}

Why do we like symmetric-information games?

Backward induction!



s^0



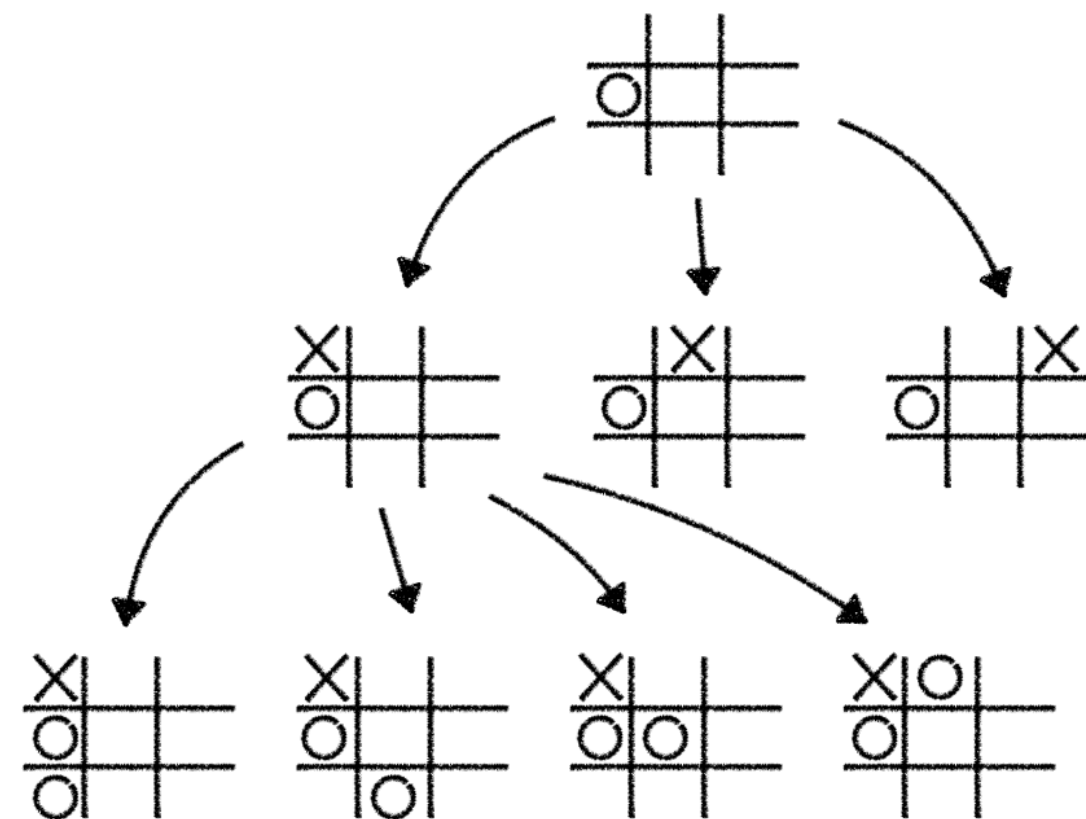
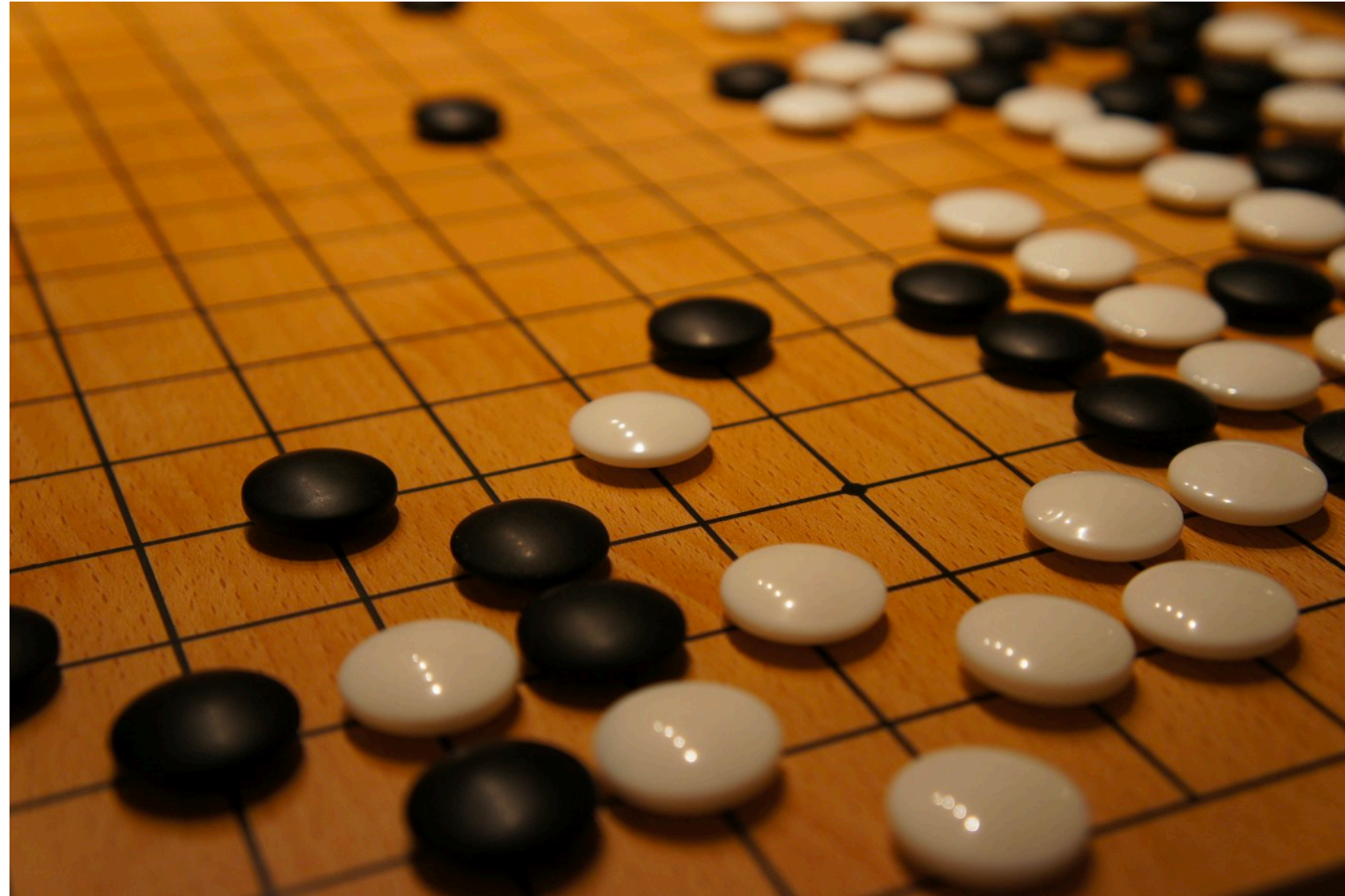
s^{T-1}

Why do we like symmetric-information games?

Search!

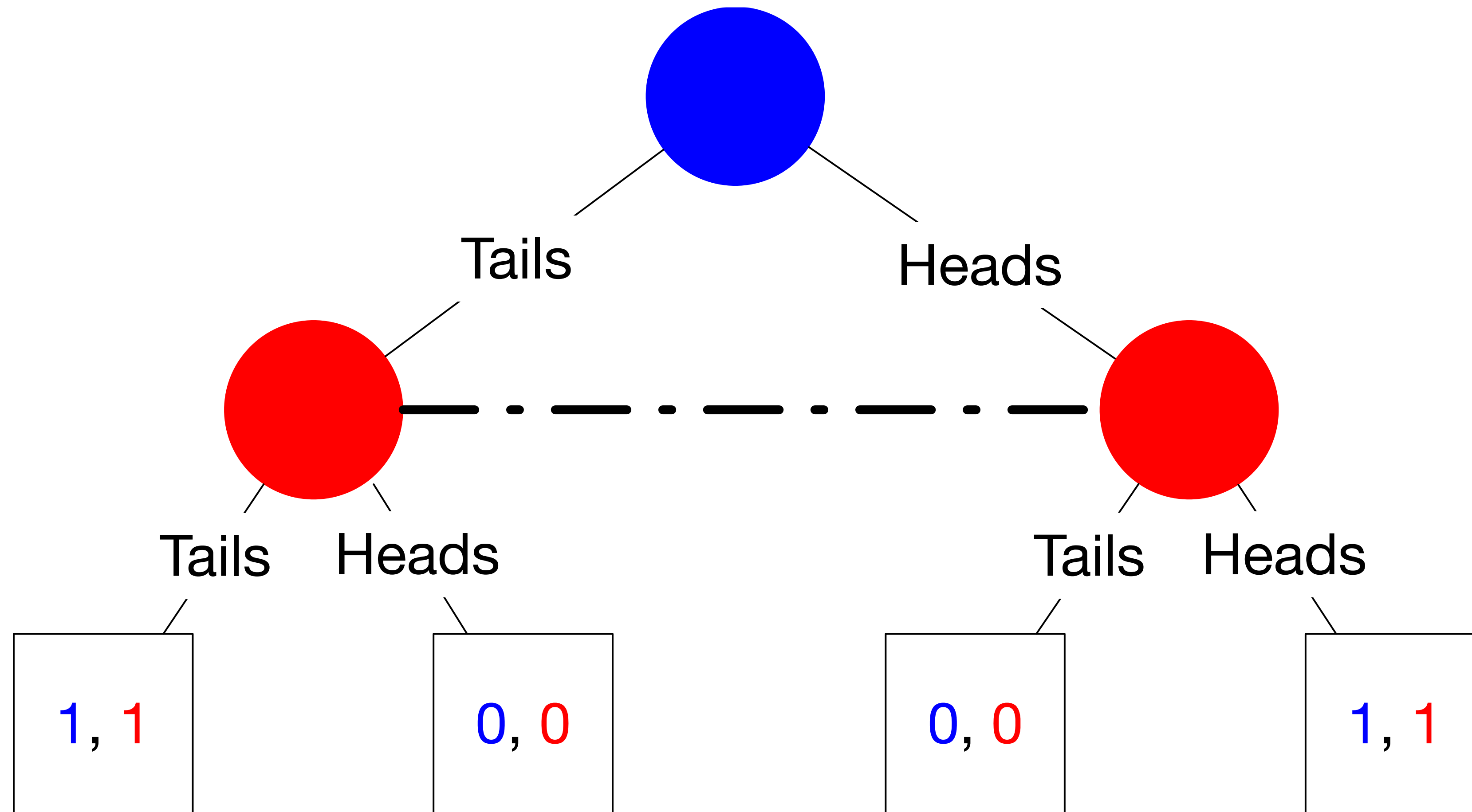
Why do we like symmetric-information games?

Search!



What happens with asymmetric information?

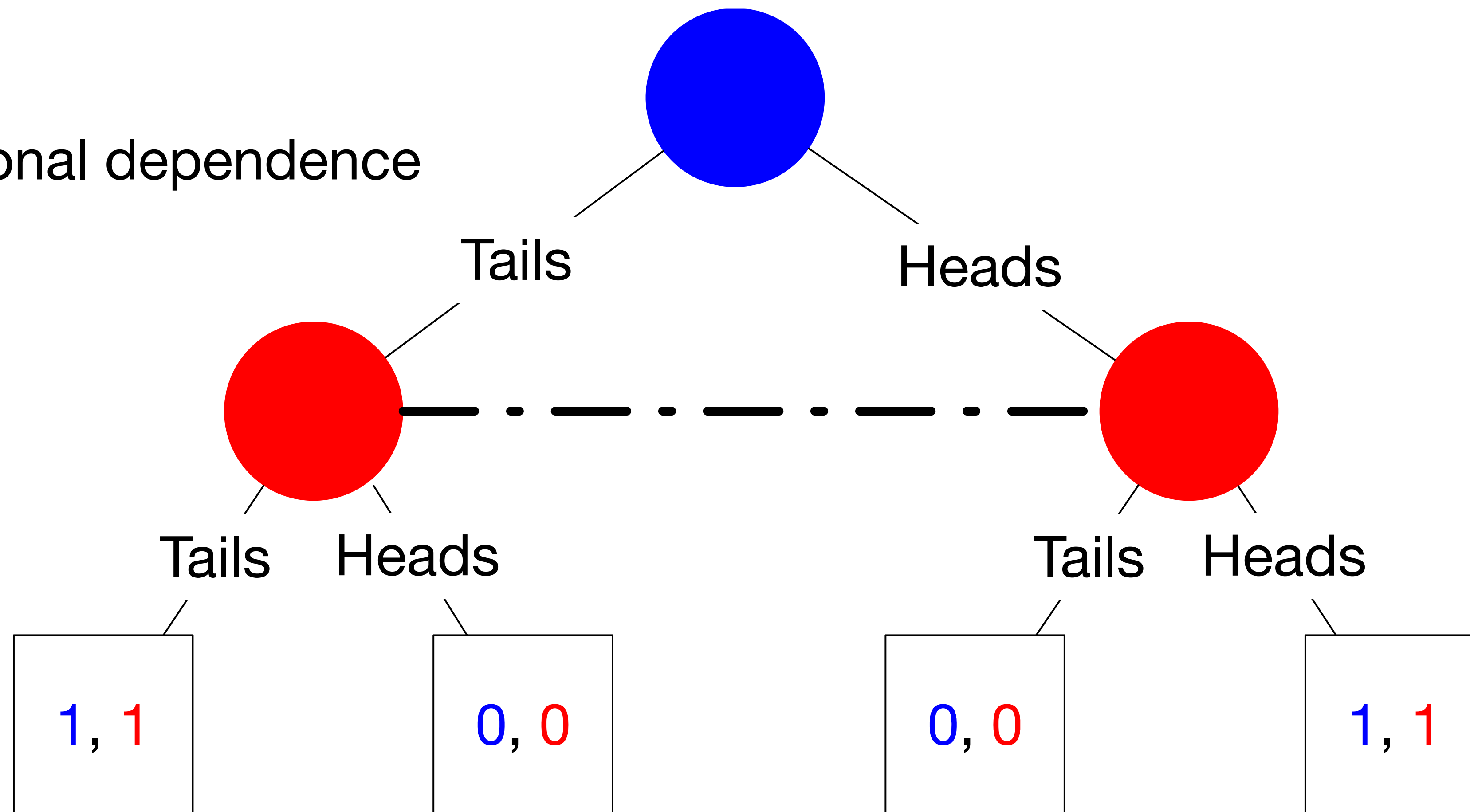
What happens with asymmetric information?



What happens with asymmetric information?

Problem:

Bidirectional dependence



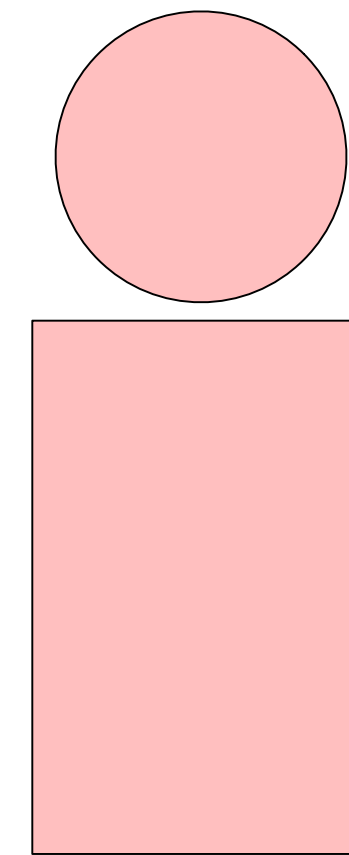
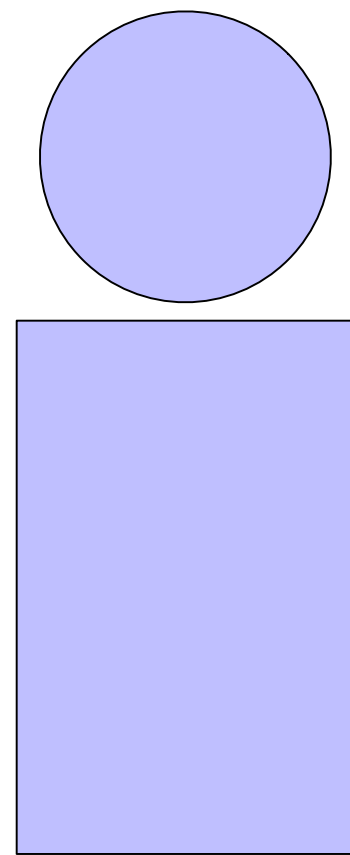
Can we get rid of information asymmetry?

Can we get rid of information asymmetry?

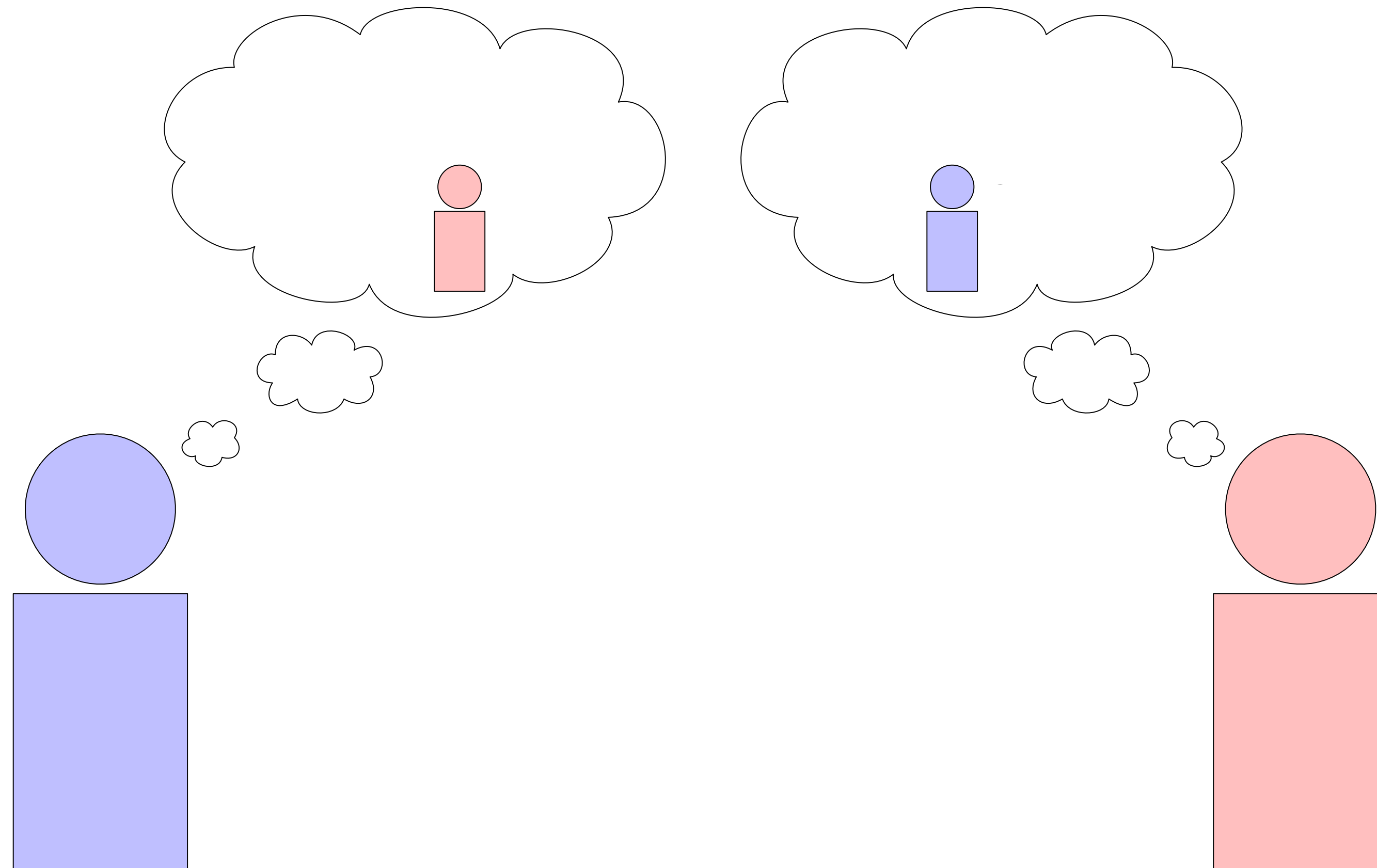
Answer: Yes, using common knowledge

What is common knowledge?

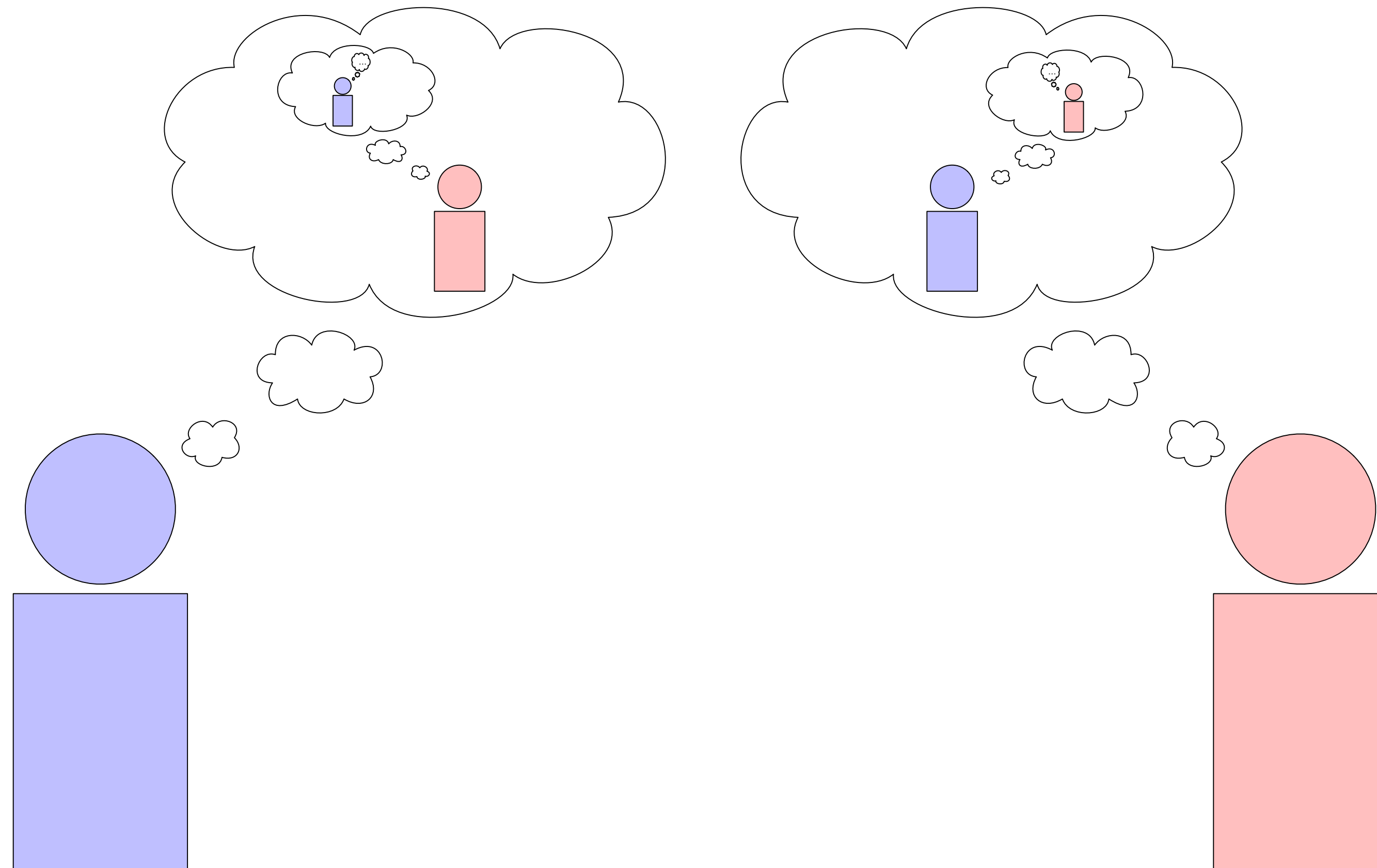
What is common knowledge?



What is common knowledge?



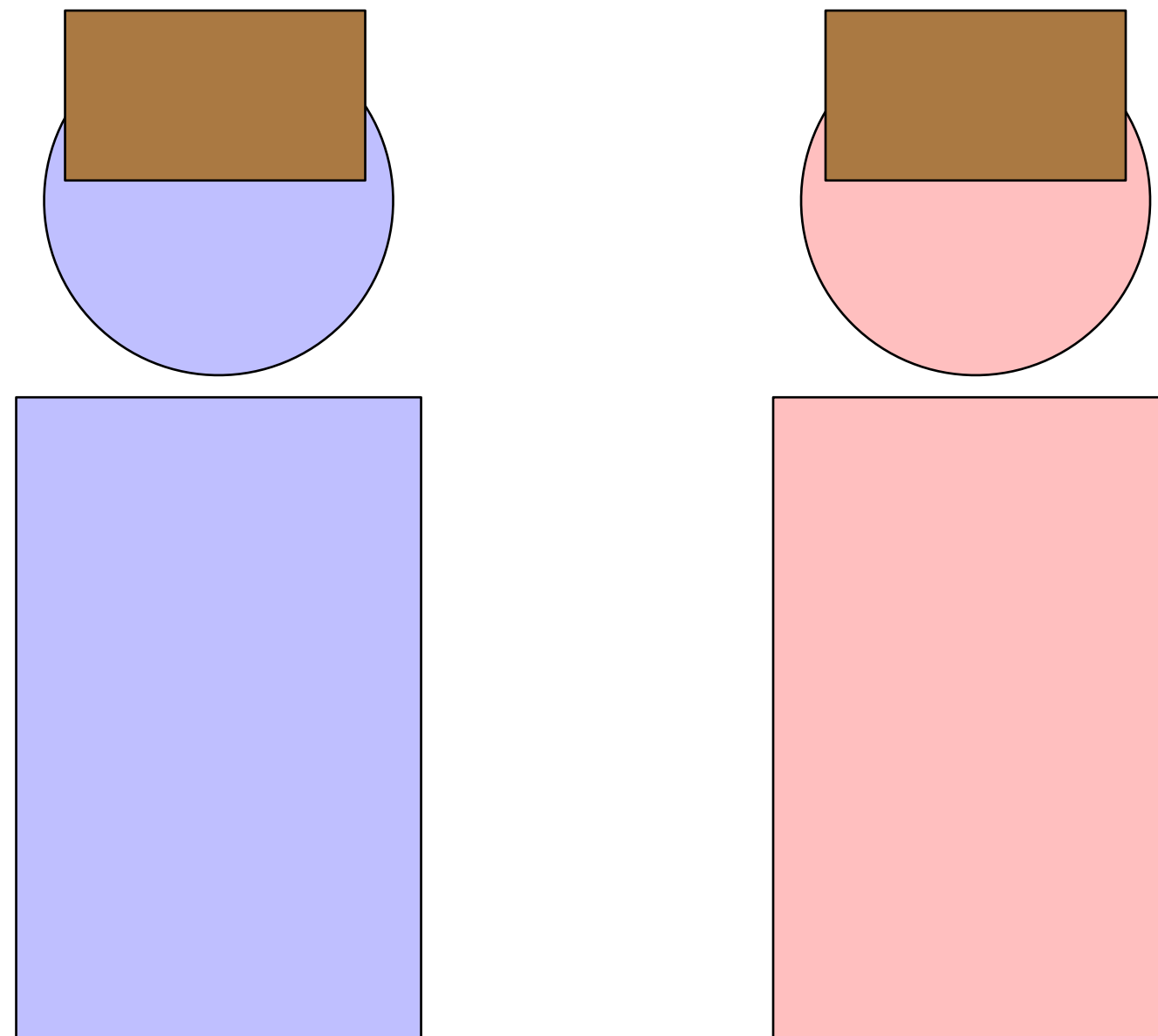
What is common knowledge?



Why is common knowledge important?

Why is common knowledge important?

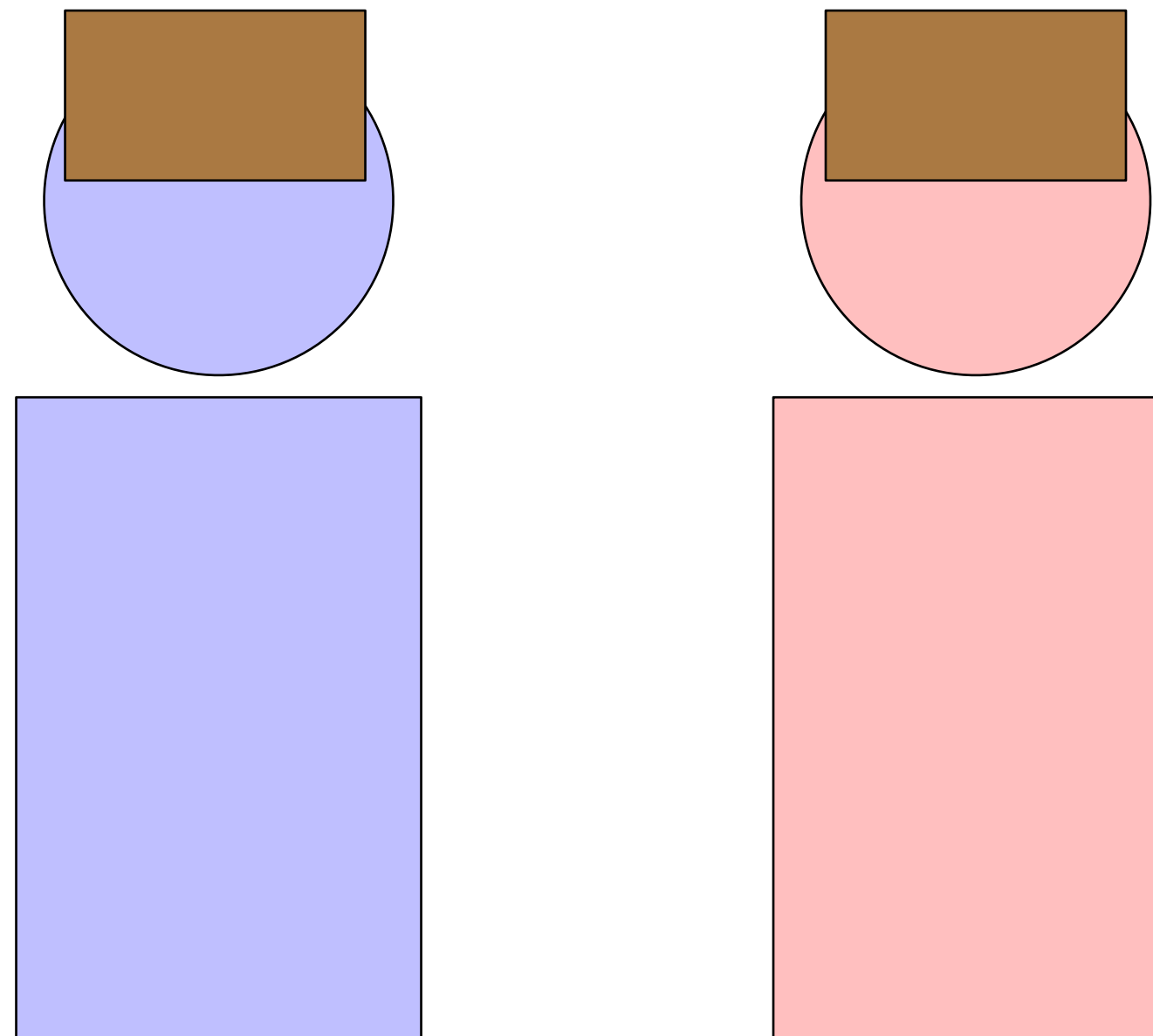
Muddy children problem:



Why is common knowledge important?

Muddy children problem:

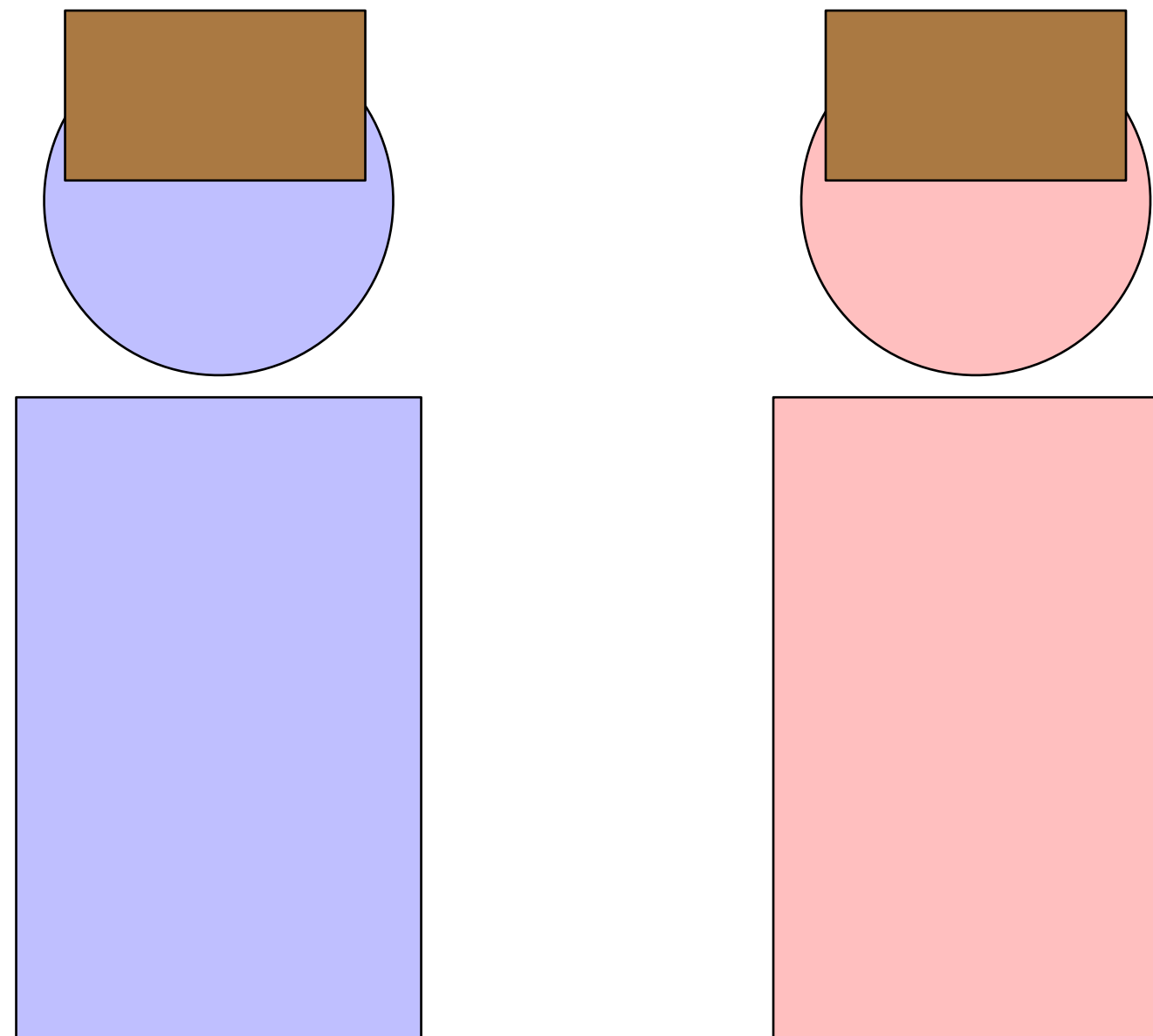
Very rational children



Why is common knowledge important?

Muddy children problem:

Very rational children

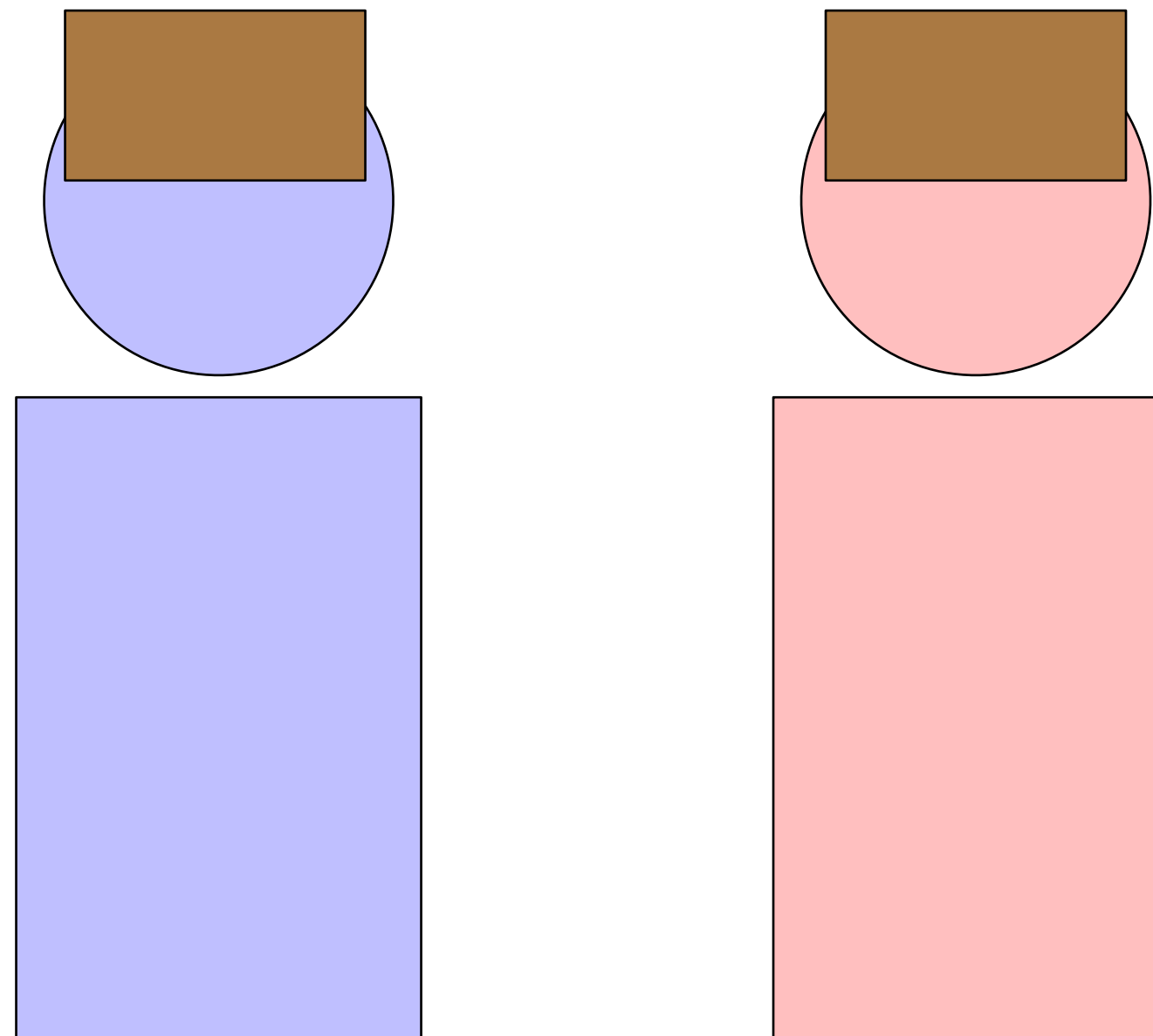


Case 1.

Why is common knowledge important?

Muddy children problem:

Very rational children

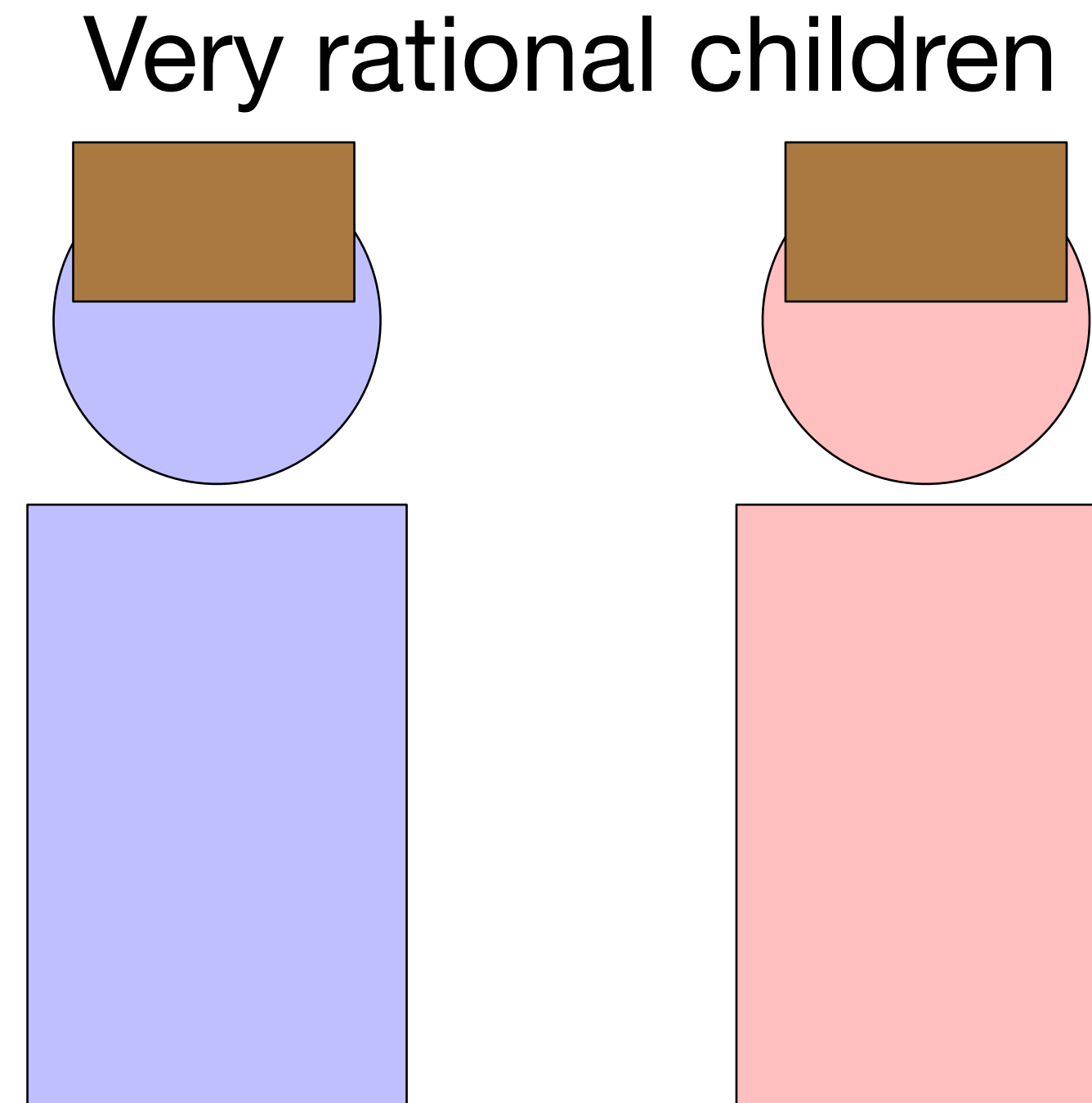


Case 1.

Father says “Raise your hand if you know you have mud on your head,” pauses for a few seconds, then repeats “Raise your hand if you know you have mud on your head.”

Why is common knowledge important?

Muddy children problem:



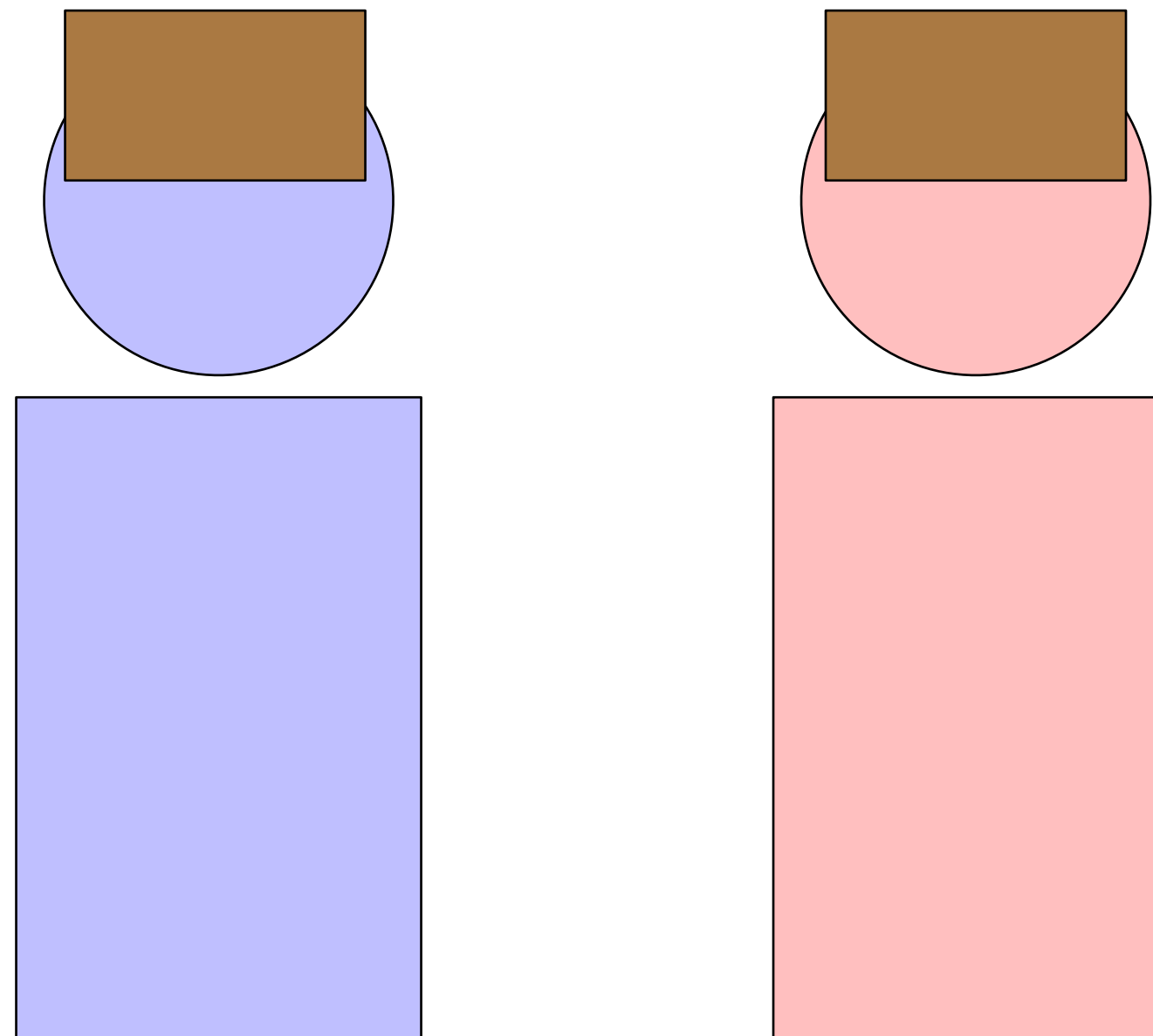
Case 2.

Father says “Raise your hand if you know you have mud on your head,” pauses for a few seconds, then repeats “Raise your hand if you know you have mud on your head.”

Why is common knowledge important?

Muddy children problem:

Very rational children



Case 2.

Father says “Raise your hand if you know you have mud on your head,” pauses for a few seconds, then repeats “Raise your hand if you know you have mud on your head.”

At least one
of you has
mud on your
head

How can we use common knowledge?

How can we use common knowledge?

Decentralized Stochastic Control with Partial History Sharing: A Common Information Approach

Ashutosh Nayyar, Aditya Mahajan and Demosthenis Teneketzis

How can we use common knowledge?

Decentralized Stochastic Control with Partial History Sharing: A Common Information Approach

Ashutosh Nayyar, Aditya Mahajan and Demosthenis Teneketzis

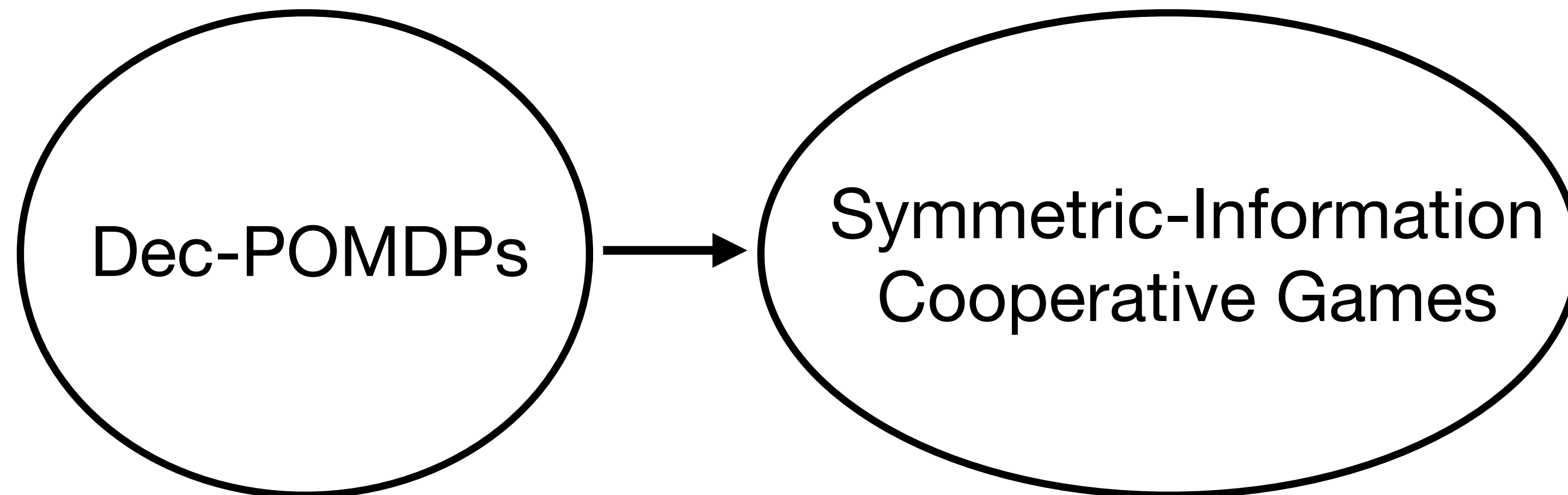


Dec-POMDPs

How can we use common knowledge?

Decentralized Stochastic Control with Partial History Sharing: A Common Information Approach

Ashutosh Nayyar, Aditya Mahajan and Demosthenis Teneketzis



How can we use common knowledge?

Decentralized Stochastic Control with Partial History Sharing: A Common Information Approach

Ashutosh Nayyar, Aditya Mahajan and Demosthenis Teneketzis



Nayyar's reduction

Nayyar's reduction

1. Player's observe public observations

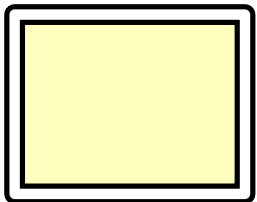

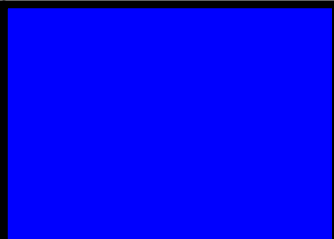










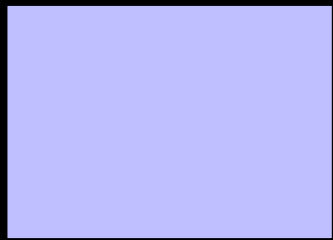
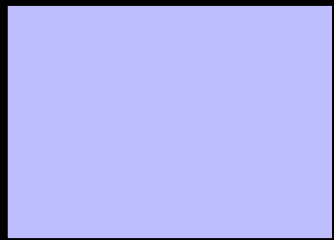
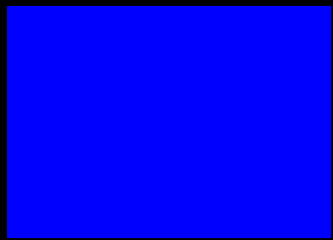
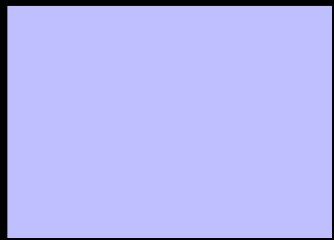
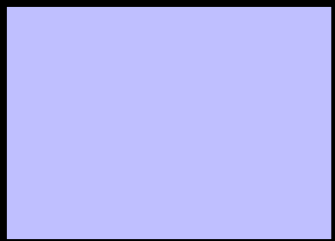
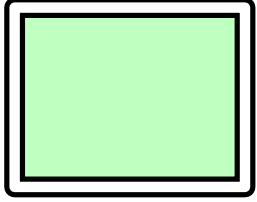
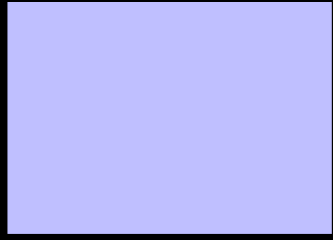
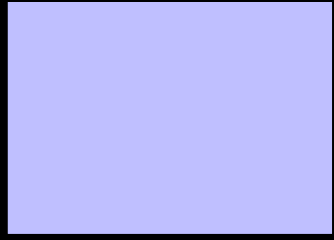
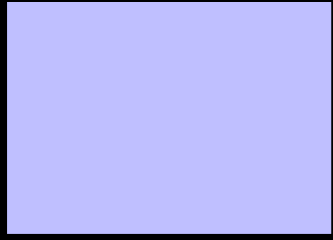
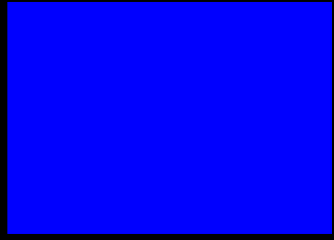
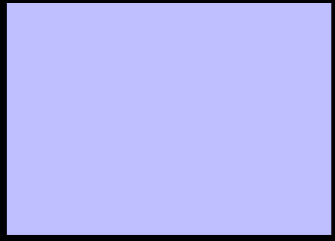
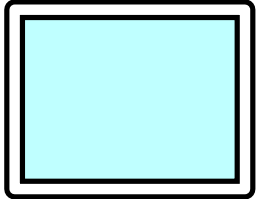
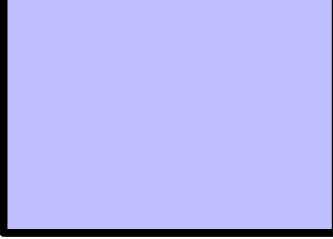

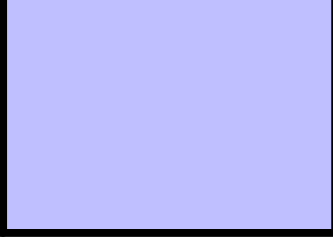
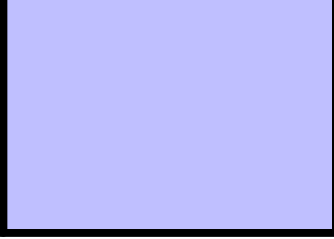
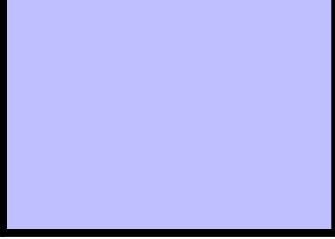
Nayyar's reduction

1. Player's observe public observations (what is public knowledge?)

Nayyar's reduction

1. Player's observe public observations
2. Players act using prescriptions

Information State

		Action				
		a	b	c	d	e
Information State						
						
						
						
						

Nayyar's reduction

1. Player's observe public observations
2. Players act using prescriptions

Claim. This is a symmetric-information cooperative game and solving it solves the original Dec-POMDP.

Nayyar's reduction

1. Player's observe public observations
2. Players act using prescriptions

Claim. This is a symmetric-information cooperative game and solving it solves the original Dec-POMDP.

Why?

So what's a public belief state?

So what's a public belief state?

Def. A public belief state is a posterior over histories given:

So what's a public belief state?

Def. A public belief state is a posterior over histories given:

1. The joint policy

So what's a public belief state?

Def. A public belief state is a posterior over histories given:

1. The joint policy
2. Public information

So what's a public belief state?

Def. A public belief state is a posterior over histories given:

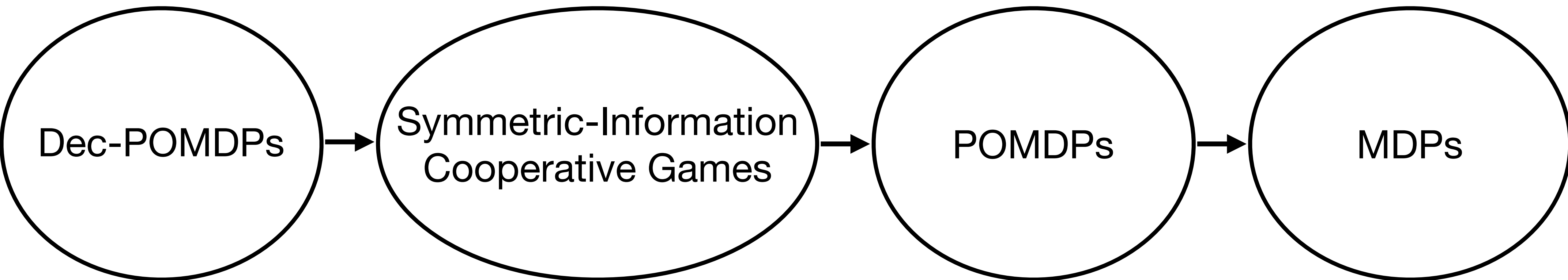
1. The joint policy
2. Public information



So what's a public belief state?

Def. A public belief state is a posterior over histories given:

1. The joint policy
2. Public information



How can we leverage this reduction?

How can we leverage this reduction?

Answer 1: Solving Dec-POMDPs with backward induction

Optimally Solving Dec-POMDPs as Continuous-State MDPs

Jilles Steeve Dibangoye

Univ de Lyon

INSA-Lyon, CITI-Inria, F-69621, France

JILLES-STEEVE.DIBANGOYE@INSA-LYON.FR

Christopher Amato

University of New Hampshire

Durham, NH, USA

CAMATO@CS.UNH.EDU

Olivier Buffet

François Charpillet

Inria — Université de Lorraine — CNRS

Villers-lès-Nancy, F-54600, France

OLIVIER.BUFFET@INRIA.FR

FRANCOIS.CHARPILLET@INRIA.FR

How can we leverage this reduction?

Answer 2: Search

Improving Policies via Search in Cooperative Partially Observable Games

Adam Lerer

Facebook AI Research
alerer@fb.com

Hengyuan Hu

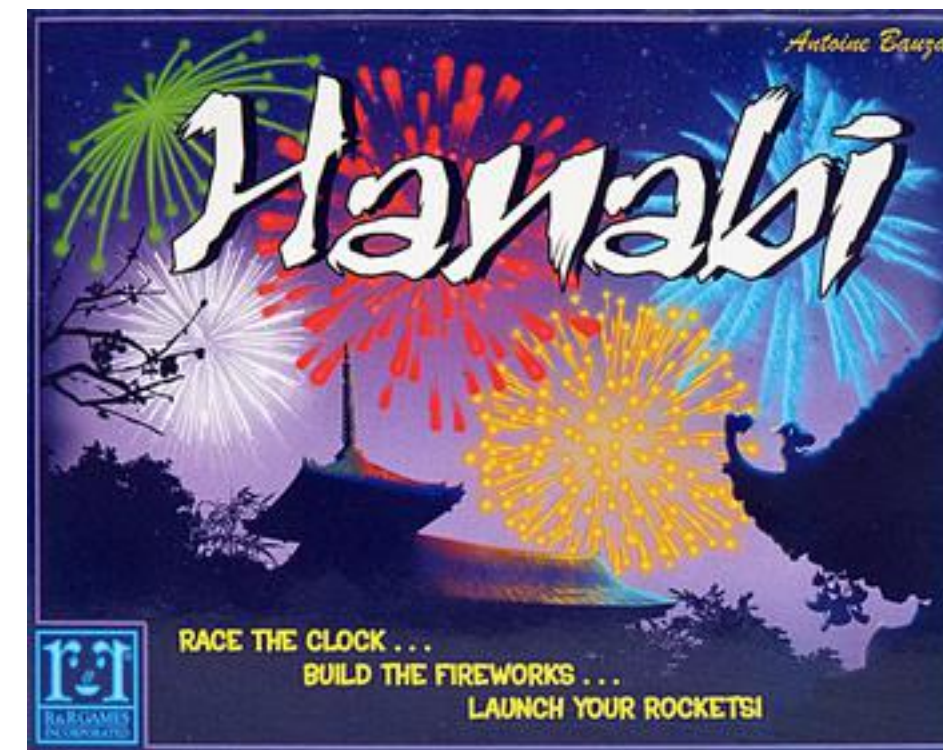
Facebook AI Research
hengyuan@fb.com

Jakob Foerster

Facebook AI Research
jnf@fb.com

Noam Brown

Facebook AI Research
noambrown@fb.com



Reductions for other classes of games?

Reductions for other classes of games?

Answer: Yes

Public Information Representation for Adversarial Team Games

Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Piazza Leonardo da Vinci, 32, 20133, Milano, Italy
luca5.carminati@mail.polimi.it, {federico.cacciamani, marco.ciccone, nicola.gatti}@polimi.it

Team Correlated Equilibria in Zero-Sum Extensive-Form Games via Tree Decompositions

Brian Hu Zhang,¹ Tuomas Sandholm^{1,2,3,4}

¹Computer Science Department, Carnegie Mellon University

²Strategic Machine, Inc.

³Strategy Robot, Inc.

⁴Optimized Markets, Inc.

{bh Zhang, sandholm}@cs.cmu.edu

Reductions for other classes of games?

Answer: Yes

Public Information Representation for Adversarial Team Games

Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Piazza Leonardo da Vinci, 32, 20133, Milano, Italy
luca5.carminati@mail.polimi.it, {federico.cacciamani, marco.ciccone, nicola.gatti}@polimi.it



Two-Team Zero-
Sum Games

Team Correlated Equilibria in Zero-Sum Extensive-Form Games via Tree Decompositions

Brian Hu Zhang,¹ Tuomas Sandholm^{1,2,3,4}

¹Computer Science Department, Carnegie Mellon University

²Strategic Machine, Inc.

³Strategy Robot, Inc.

⁴Optimized Markets, Inc.

{bh Zhang, sandholm}@cs.cmu.edu

Reductions for other classes of games?

Answer: Yes

Public Information Representation for Adversarial Team Games

Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Piazza Leonardo da Vinci, 32, 20133, Milano, Italy
luca5.carminati@mail.polimi.it, {federico.cacciamani, marco.ciccone, nicola.gatti}@polimi.it

Team Correlated Equilibria in Zero-Sum Extensive-Form Games via Tree Decompositions

Brian Hu Zhang,¹ Tuomas Sandholm^{1,2,3,4}

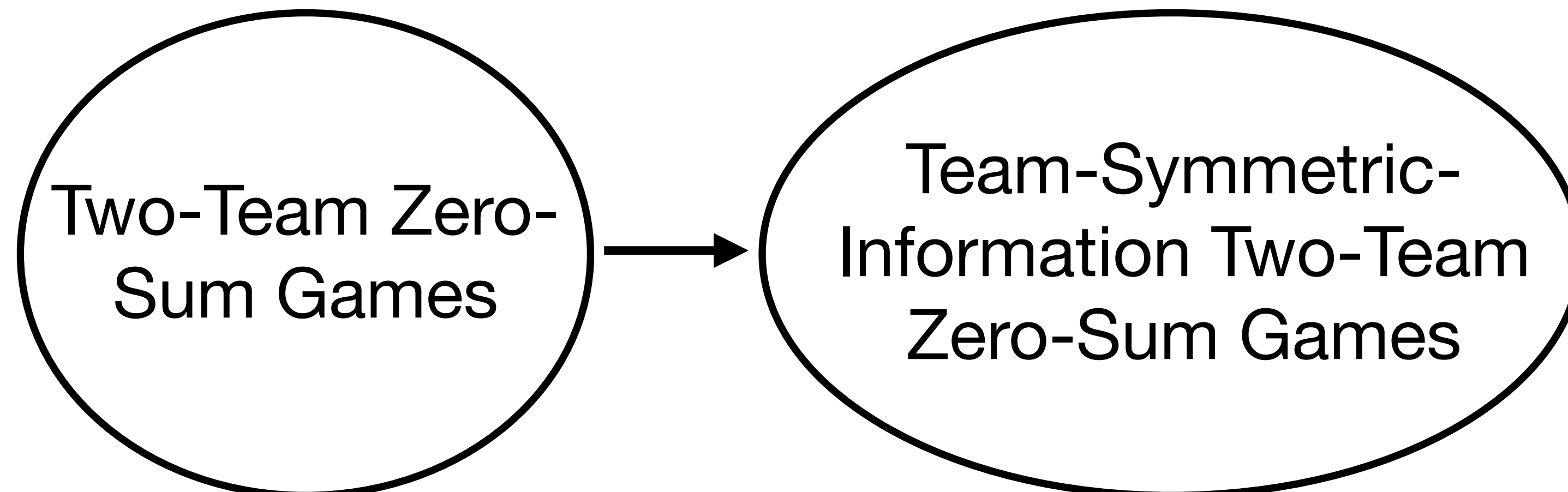
¹Computer Science Department, Carnegie Mellon University

²Strategic Machine, Inc.

³Strategy Robot, Inc.

⁴Optimized Markets, Inc.

{bh Zhang, sandholm}@cs.cmu.edu



Reductions for other classes of games?

Answer: Yes

Public Information Representation for Adversarial Team Games

Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
Piazza Leonardo da Vinci, 32, 20133, Milano, Italy
luca5.carminati@mail.polimi.it, {federico.cacciamani, marco.ciccone, nicola.gatti}@polimi.it

Team Correlated Equilibria in Zero-Sum Extensive-Form Games via Tree Decompositions

Brian Hu Zhang,¹ Tuomas Sandholm^{1,2,3,4}

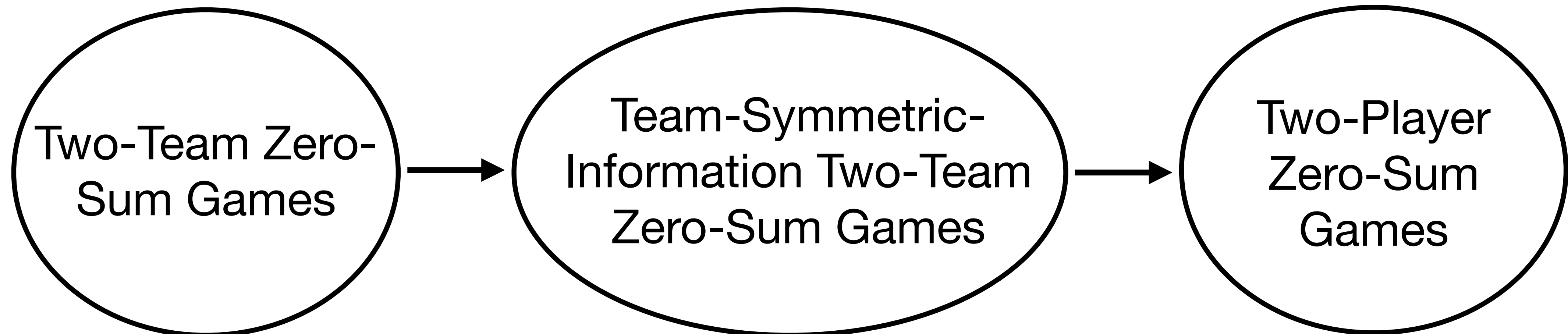
¹Computer Science Department, Carnegie Mellon University

²Strategic Machine, Inc.

³Strategy Robot, Inc.

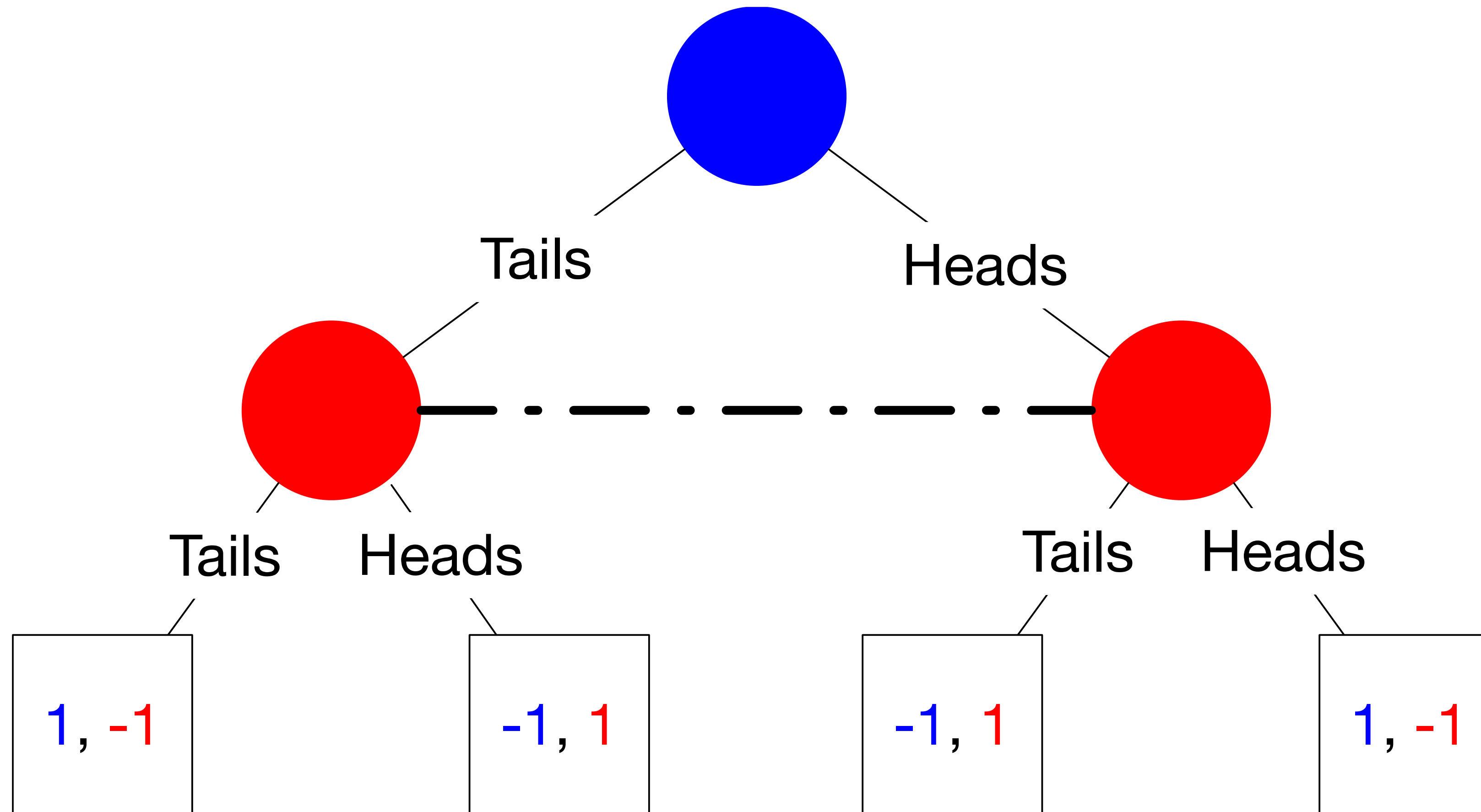
⁴Optimized Markets, Inc.

{bhzzhang, sandholm}@cs.cmu.edu

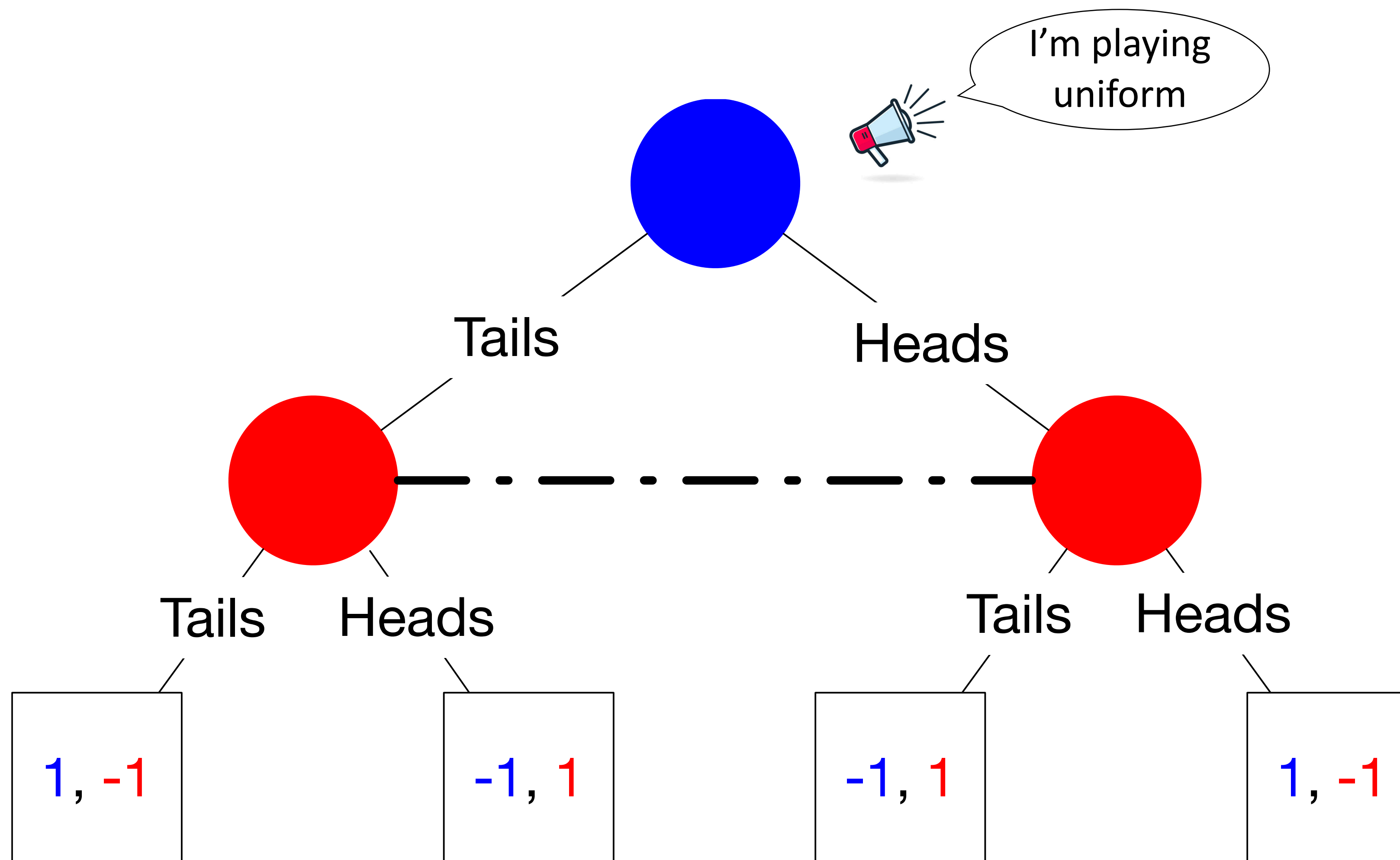


What about two-player zero-sum?

What about two-player zero-sum?

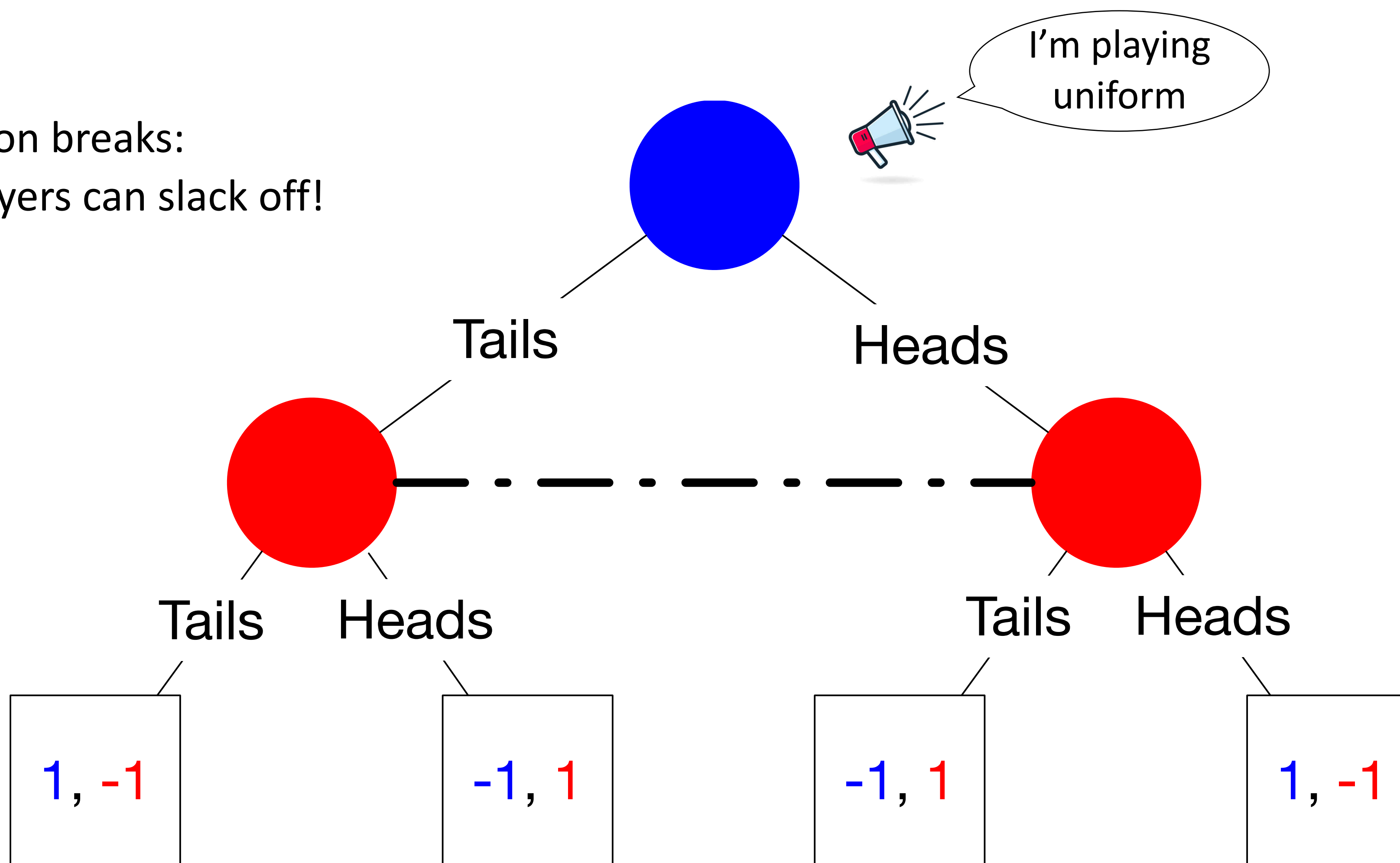


What about two-player zero-sum?



What about two-player zero-sum?

Backward induction breaks:
Latter moving players can slack off!



How can we address this?

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

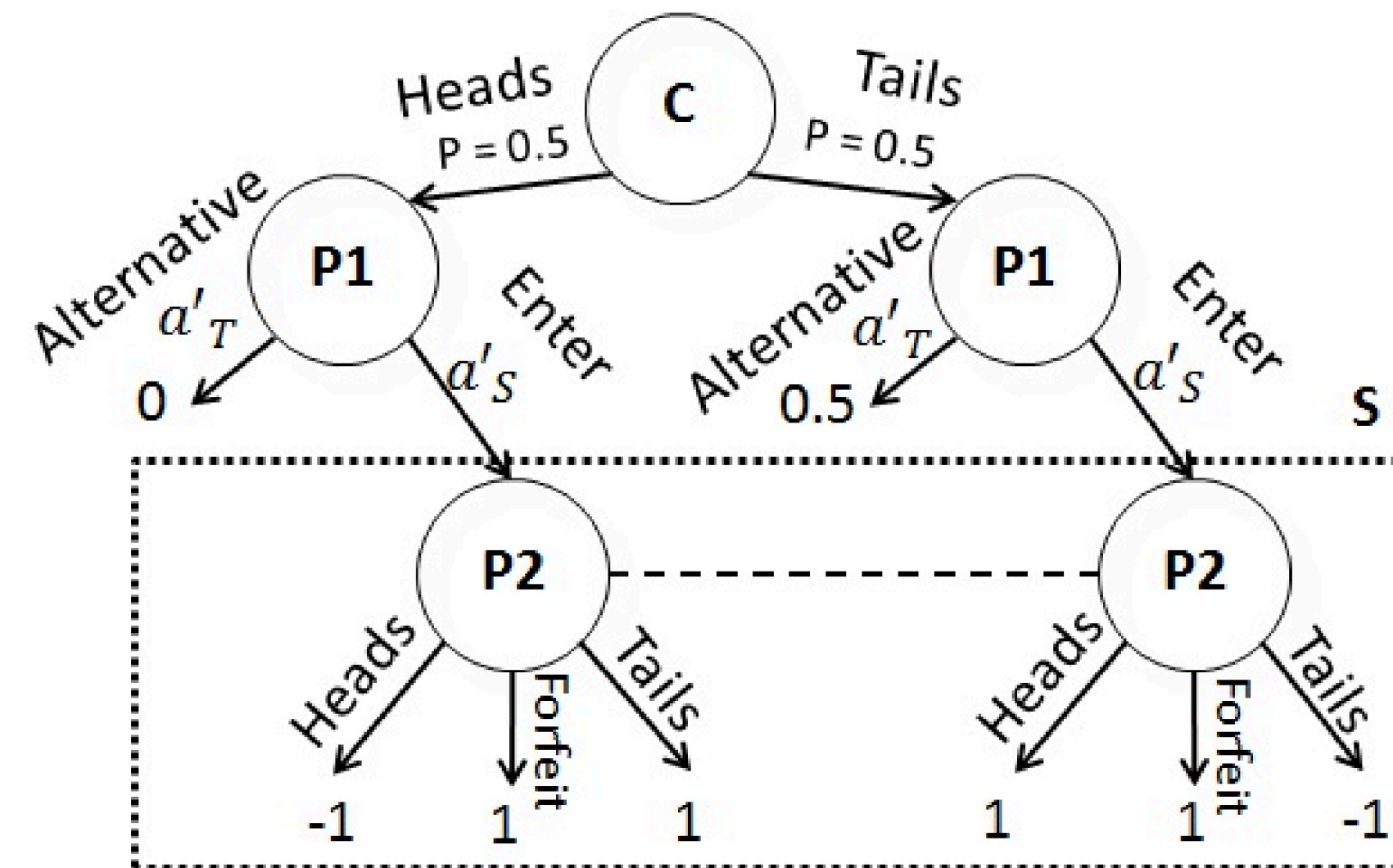
[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*



DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

2. ReBeL

Combining Deep Reinforcement Learning and Search for Imperfect-Information Games

Noam Brown* Anton Bakhtin* Adam Lerer Qucheng Gong
Facebook AI Research
{noambrown,yolo,alerer,qucheng}@fb.com

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

2. ReBeL

Combining Deep Reinforcement Learning and Search for Imperfect-Information Games

Noam Brown* Anton Bakhtin* Adam Lerer Qucheng Gong
Facebook AI Research
{noambrown,yolo,alerer,qucheng}@fb.com

3. Regularization

Abstracting Imperfect Information Away from Two-Player Zero-Sum Games

Samuel Sokota^{1†} Ryan D'Orazio² Chun Kai Ling¹ David J. Wu³ J. Zico Kolter^{1,4} Noam Brown³

How can we address this?

1. Gadget games

Superhuman AI for heads-up no-limit poker: Libratus beats top professionals

Noam Brown and Tuomas Sandholm*

DeepStack: Expert-Level Artificial Intelligence in Heads-Up No-Limit Poker

Matej Moravčík^{♠,♥,†}, Martin Schmid^{♠,♥,†}, Neil Burch[♠], Viliam Lisý^{♠,♣},
Dustin Morrill[♠], Nolan Bard[♠], Trevor Davis[♠],
Kevin Waugh[♠], Michael Johanson[♠], Michael Bowling^{♠,*}

[♠]Department of Computing Science, University of Alberta,
Edmonton, Alberta, T6G2E8, Canada

[♥]Department of Applied Mathematics, Charles University,
Prague, Czech Republic

[♣]Department of Computer Science, FEE, Czech Technical University,
Prague, Czech Republic

2. ReBeL

Combining Deep Reinforcement Learning and Search for Imperfect-Information Games

Noam Brown* Anton Bakhtin* Adam Lerer Qucheng Gong
Facebook AI Research
{noambrown,yolo,alerer,qucheng}@fb.com

3. Regularization

Abstracting Imperfect Information Away from Two-Player Zero-Sum Games

Samuel Sokota^{1†} Ryan D'Orazio² Chun Kai Ling¹ David J. Wu³ J. Zico Kolter^{1,4} Noam Brown³

What does regularization mean?

$$\max_{\pi_0} \min_{\pi_1} \mathbb{E} \left[\sum_t \mathcal{R}(S^t, A^t) \mid \pi_0, \pi_1 \right]$$

What does regularization mean?

$$\max_{\pi_0} \min_{\pi_1} \mathbb{E} \left[\sum_t \mathcal{R}(S^t, A^t) \mid \pi_0, \pi_1 \right]$$

vs.

$$\max_{\pi_0} \min_{\pi_1} \mathbb{E} \left[\sum_t \mathcal{R}(S^t, A^t) + \alpha \mathcal{H}(\pi_0(S_0^t)) - \alpha \mathcal{H}(\pi_1(S_1^t)) \mid \pi \right]$$

Why does regularization fix slacking off?

Why does regularization fix slacking off?

Claim. Regularized equilibria of the public game are regularized equilibria of the original game.

Why does regularization fix slacking off?

Claim. Regularized equilibria of the public game are regularized equilibria of the original game.

Why?

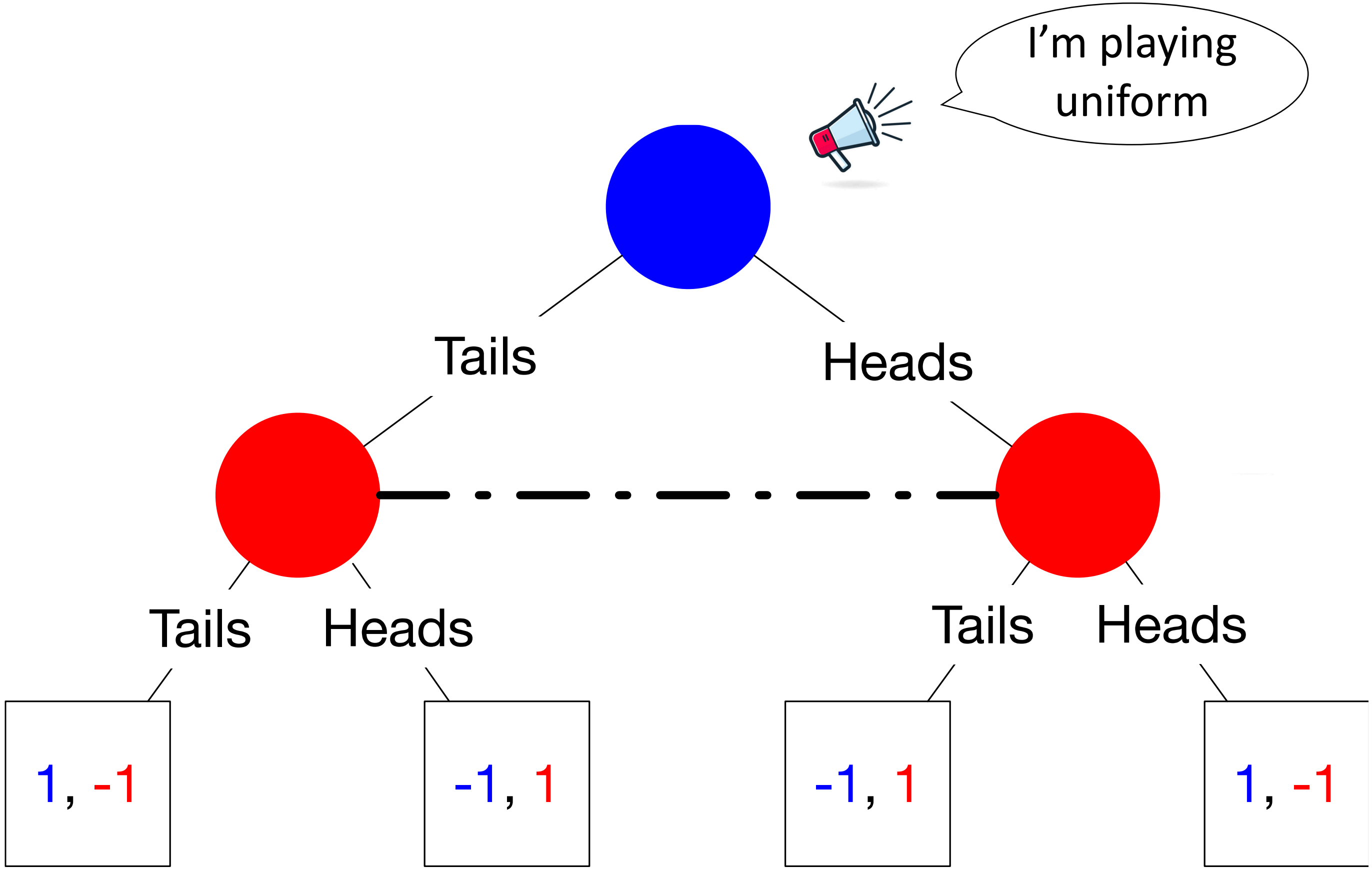
Why does regularization fix slacking off?

Claim. Regularized equilibria of the public game are regularized equilibria of the original game.

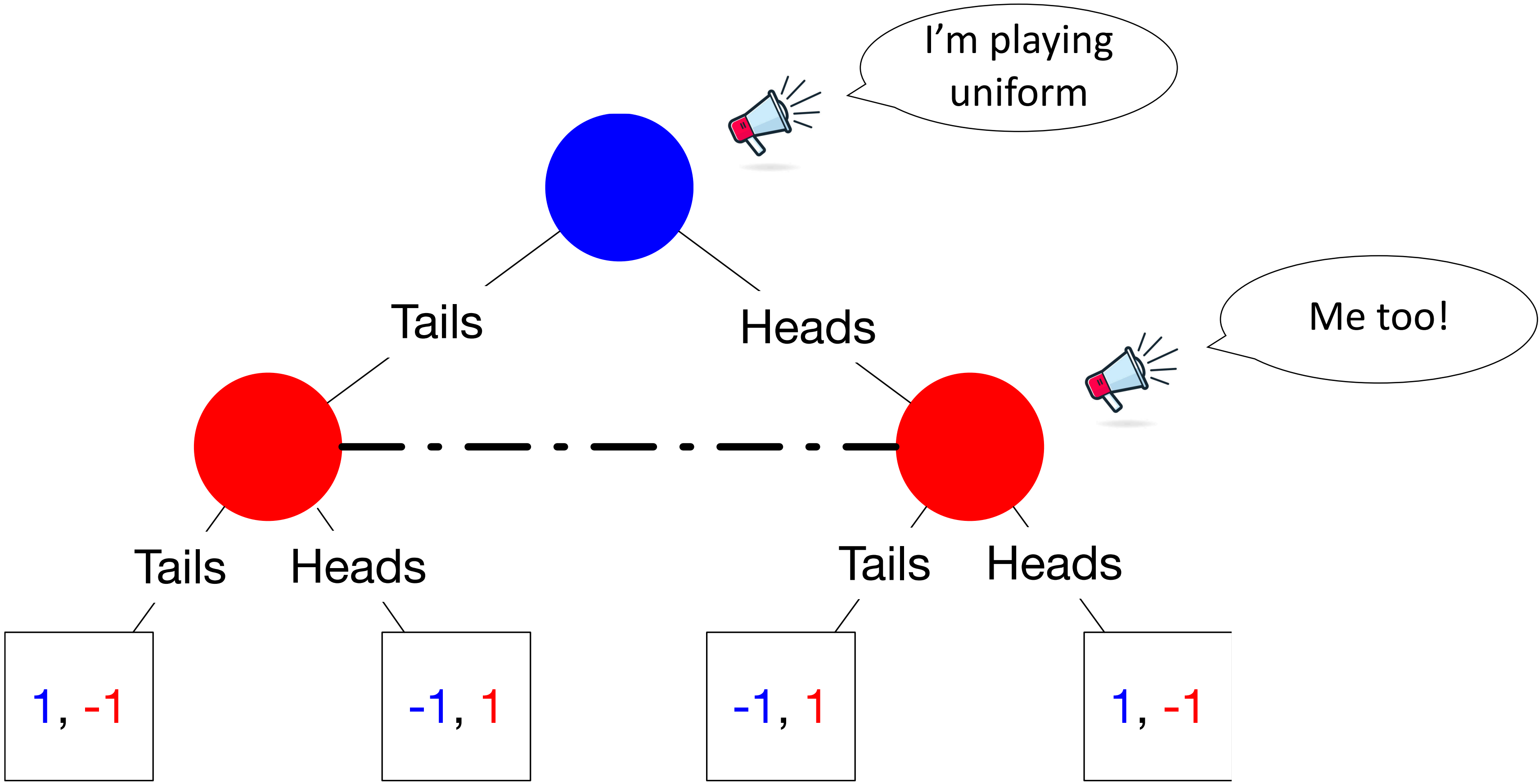
Why?

(Informal) answer: Unique best response

Why does regularization fix slacking off?



Why does regularization fix slacking off?



TLDR

TLDR

1. Common knowledge is a fundamental idea

TLDR

1. Common knowledge is a fundamental idea
2. Public belief states are the analog of Markov states

TLDR

1. Common knowledge is a fundamental idea
2. Public belief states are the analog of Markov states
3. Use cases:

TLDR

1. Common knowledge is a fundamental idea
2. Public belief states are the analog of Markov states
3. Use cases:
 - i. Solving Dec-POMDPs

TLDR

1. Common knowledge is a fundamental idea
2. Public belief states are the analog of Markov states
3. Use cases:
 - i. Solving Dec-POMDPs
 - ii. Search for cooperative games (e.g. Hanabi)

TLDR

1. Common knowledge is a fundamental idea
2. Public belief states are the analog of Markov states
3. Use cases:
 - i. Solving Dec-POMDPs
 - ii. Search for cooperative games (e.g. Hanabi)
 - iii. Search in adversarial games (e.g. poker)