Reinforcement Learning Conference

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Inception: Efficiently Computable Misinformation Attacks on Markov Games

Jeremy McMahan, Young Wu, Yudong Chen, Xiaojin Zhu, and Qiaomin Xie

Motivation

Motivation

- Lying about rewards can improve outcomes.

Motivation

- Lying about rewards can improve outcomes.
- Lies can come from misinformation online.

What happens when assumptions are violated?

Example

True Game

Example

True Game

	L	R
U	$0,\ 5$	$1,\ 0$
D	$1,\ \epsilon$	$0,\ 0$
S	$1,\ 0$	$0,\ \epsilon$

Example

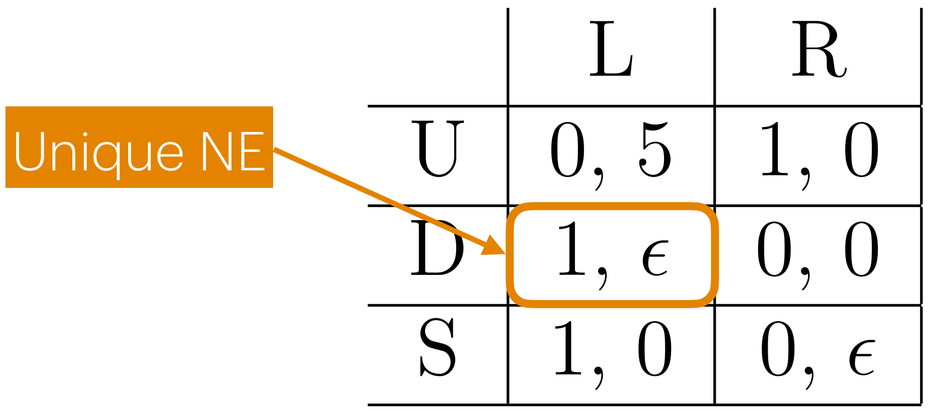
True Game

Unique NE

	L	R
U	$0,\ 5$	$1,\ 0$
D	$1,\ \epsilon$	$0,\ 0$
S	$1,\ 0$	$0,\ \epsilon$

Example

True Game



	L	R
U	$0,\ 5$	$1,\ 0$
D	$1,\ \epsilon$	$0,\ 0$
S	$1,\ 0$	$0,\ \epsilon$

If P1 is rational, it must play D, so P2 gets 0!

Example

True Game

P2 wants

Unique NE

	L	R
U	0, 5	1, 0
D	1, ϵ	0, 0
S	1, 0	0, ϵ

If P1 is rational, it must play D, so P2 gets 0!

Example

Faked Game

Example

Faked Game

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

Example

Faked Game

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

Increased

Example

Faked Game

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

Unique NE

Example

Faked Game

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

Unique NE

P1 must play U, so P2 can get 5 in true game!

Example

Faked Game

P2 wins!

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

Unique NE

P1 must play U, so P2 can get 5 in true game!

Example

Faked Game

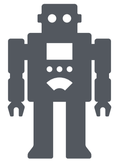
“Inception Attack”

Inception attacks are powerful,
but can they be computed?

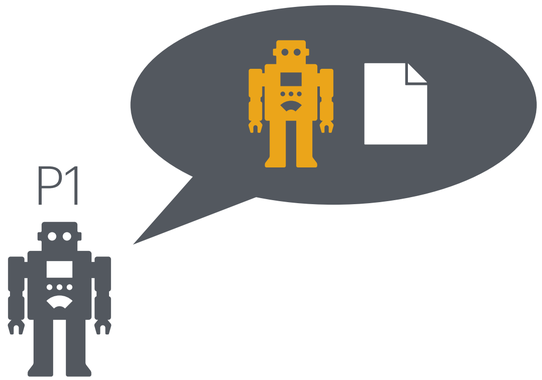
Inception Approach

Inception Approach

P1

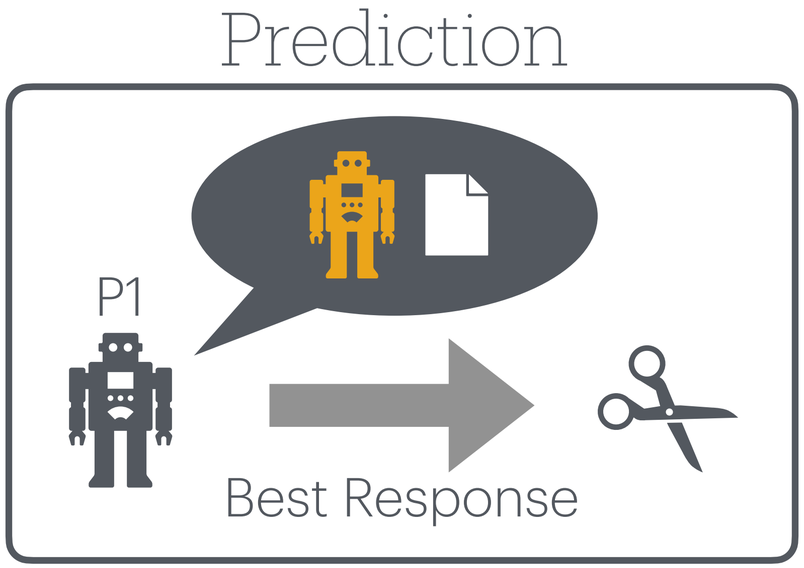


Inception Approach

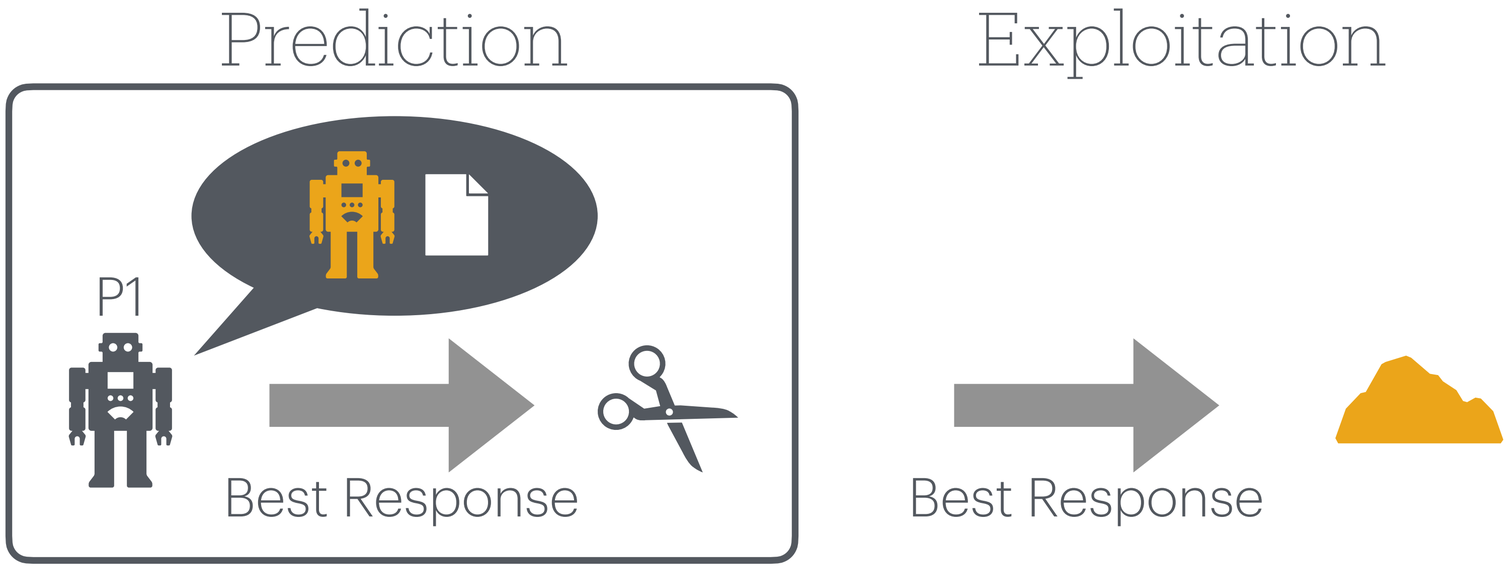


Inception Approach

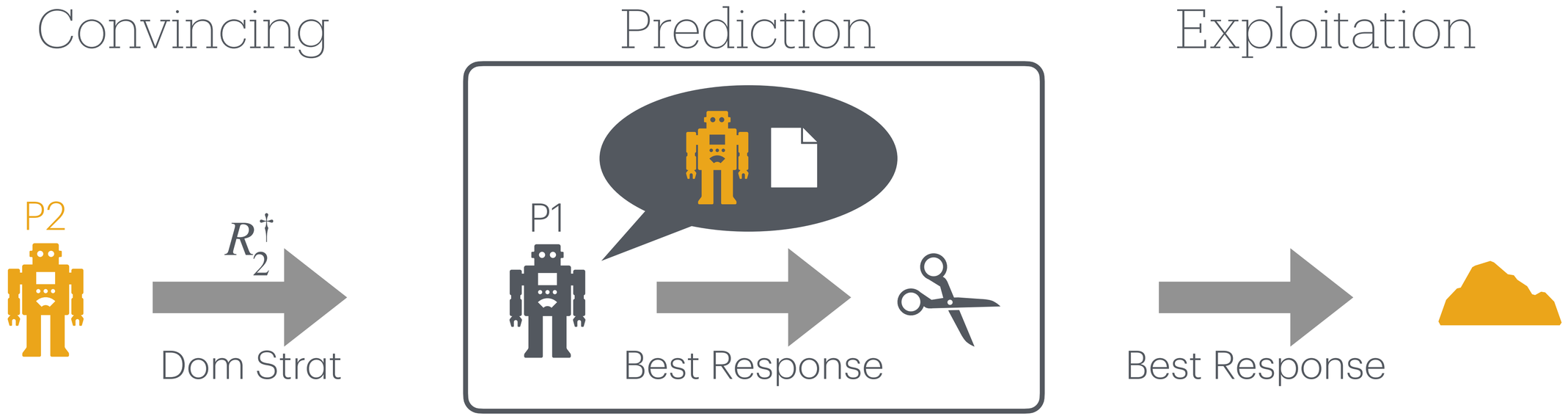
Inception Approach



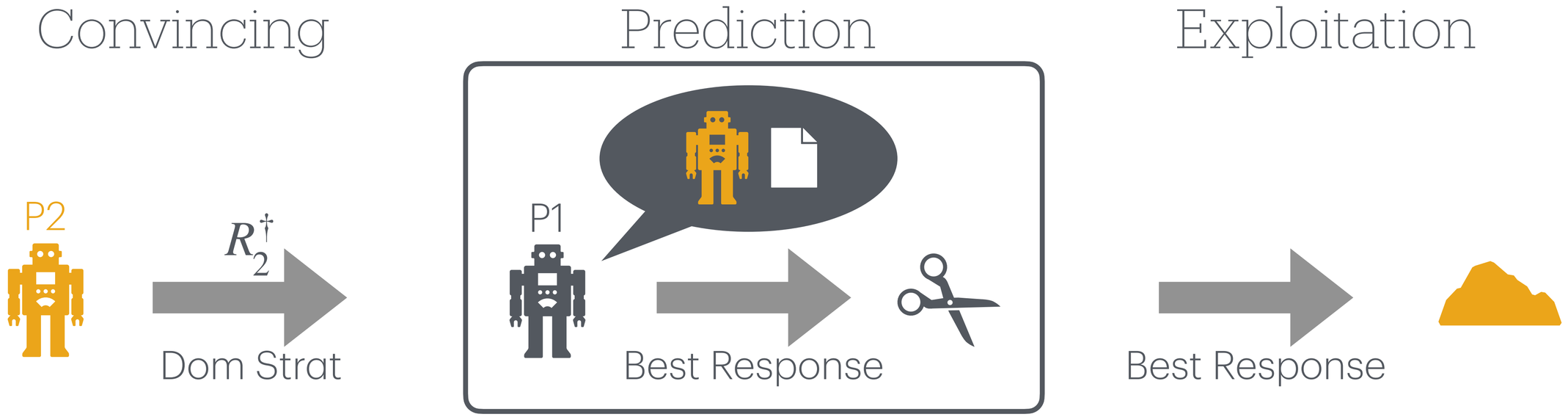
Inception Approach



Inception Approach



Inception Approach



Repeat to find the best pure strategy inception!

Algorithm

P2 fakes R

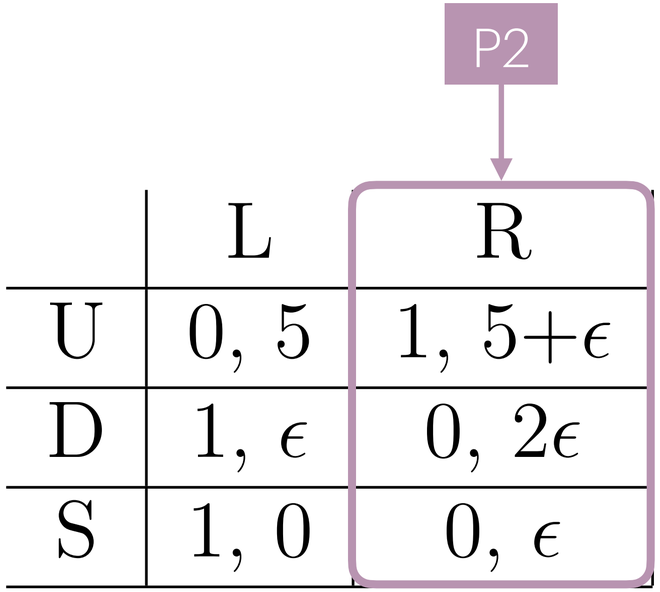
Algorithm

P2 fakes R

	L	R
U	$0,\ 5$	$1,\ 5+\epsilon$
D	$1,\ \epsilon$	$0,\ 2\epsilon$
S	$1,\ 0$	$0,\ \epsilon$

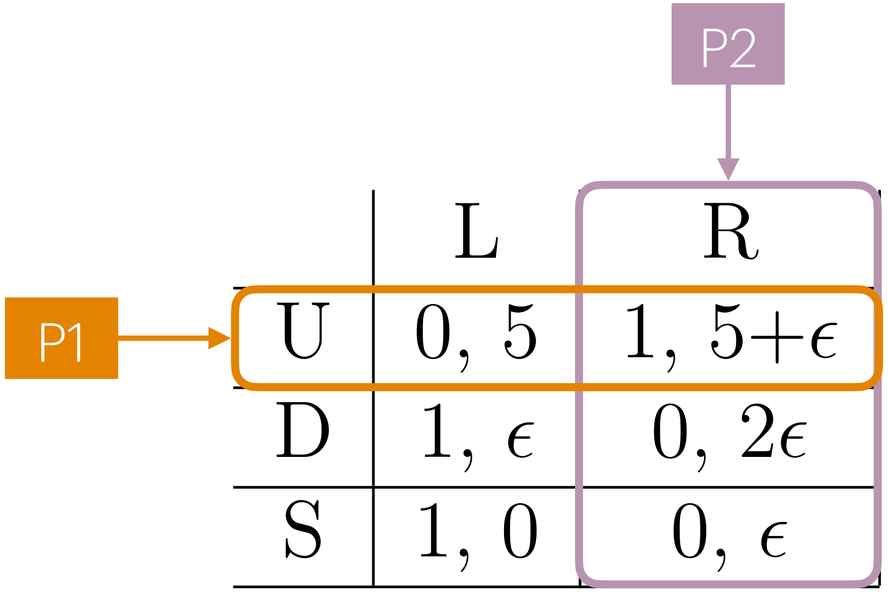
Algorithm

P2 fakes R



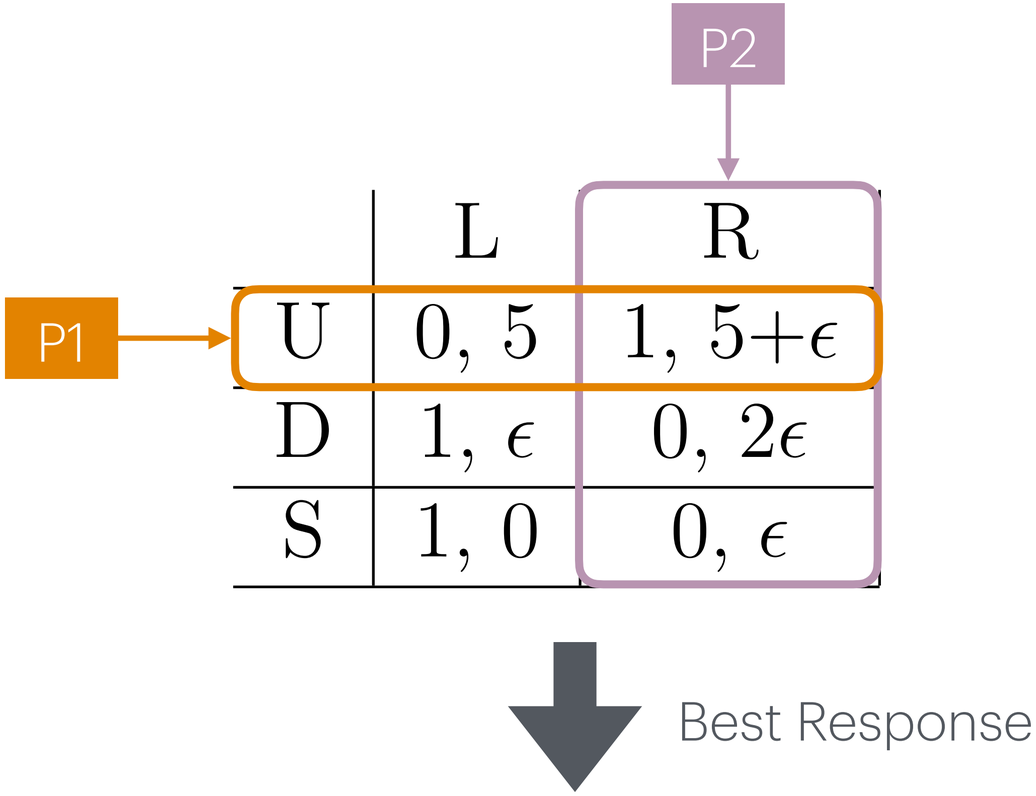
Algorithm

P2 fakes R



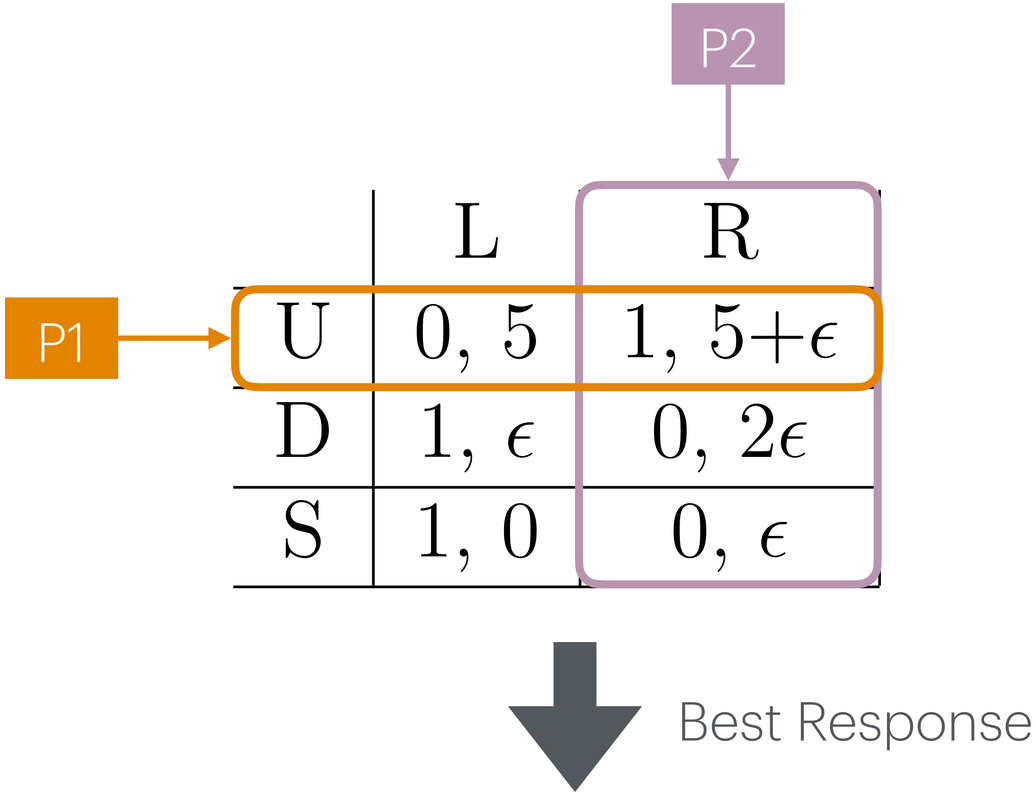
Algorithm

P2 fakes R



Algorithm

P2 fakes R



Algorithm

P2 fakes L

Algorithm

P2 fakes L

	L	R
U	$0, 5$	$1, 0$
D	$1, \epsilon$	$0, 0$
S	$1, 2\epsilon$	$0, \epsilon$

Algorithm

P2 fakes L

	L	R
U	$0, 5$	$1, 0$
D	$1, \epsilon$	$0, 0$
S	$1, 2\epsilon$	$0, \epsilon$

Increased

Algorithm

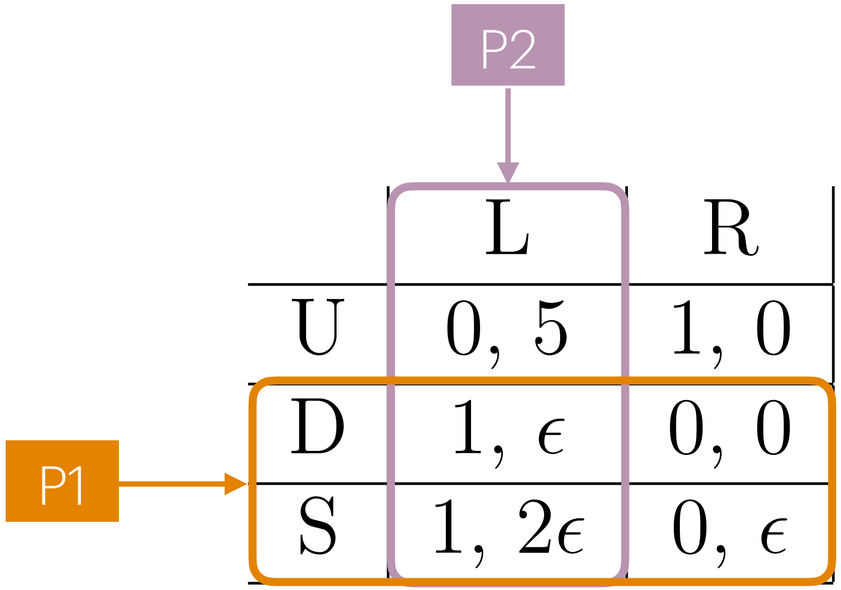
P2 fakes L

P2

	L	R
U	$0, 5$	$1, 0$
D	$1, \epsilon$	$0, 0$
S	$1, 2\epsilon$	$0, \epsilon$

Algorithm

P2 fakes L



Algorithm

P2 fakes L

P2

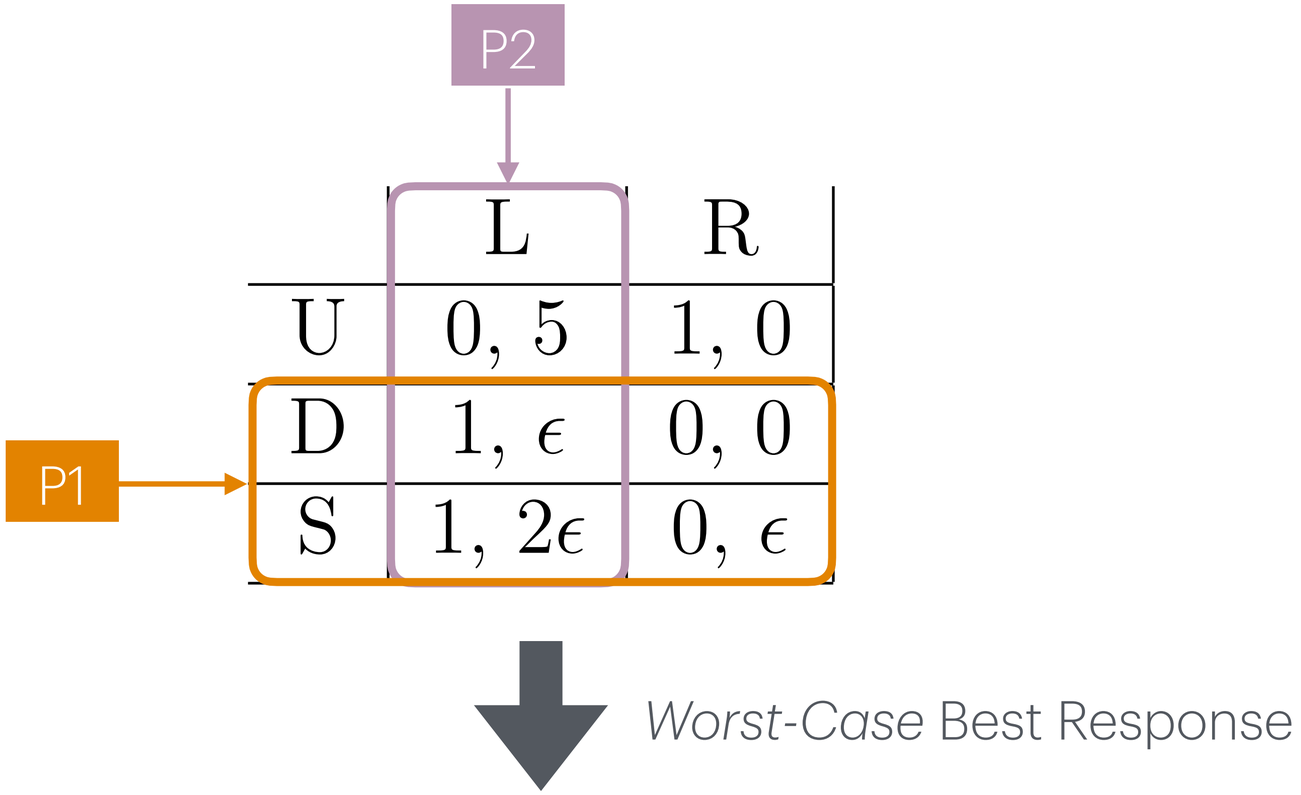
	L	R
U	0, 5	1, 0
D	1, ϵ	0, 0
S	1, 2ϵ	0, ϵ

P1

Tie!

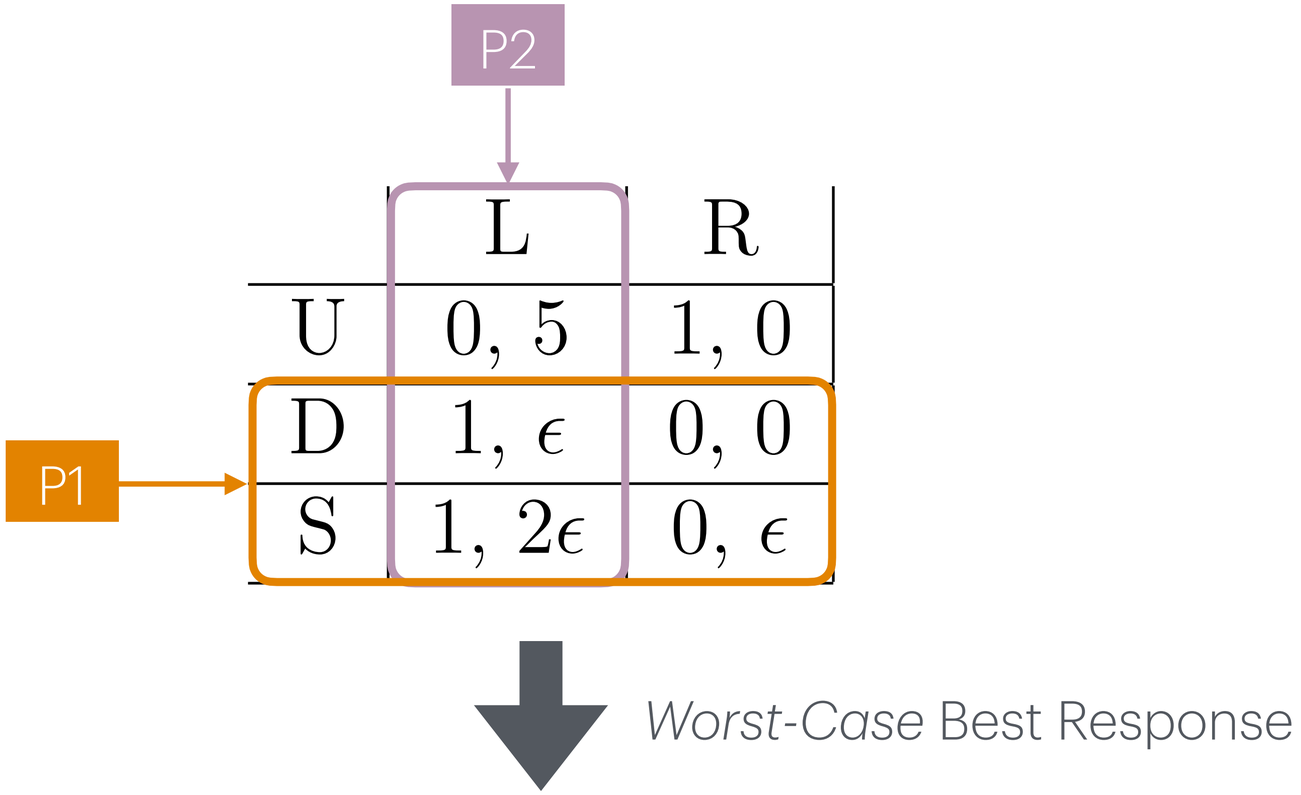
Algorithm

P2 fakes L



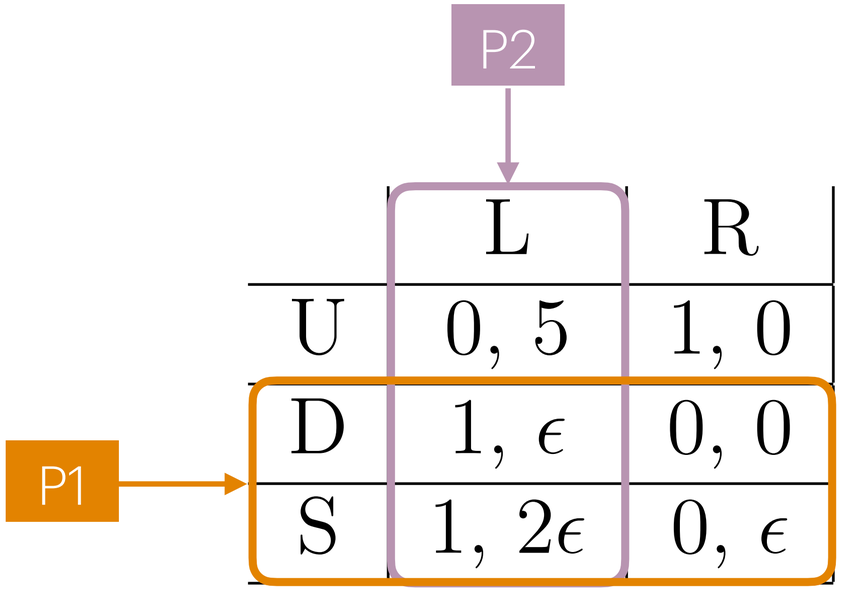
Algorithm

P2 fakes L



Algorithm

P2 fakes L



Worst-Case Best Response

Solvable by an LP!

P2 gets 2ϵ

Conclusion

Conclusion

Misinformation attacks can be computed in polynomial time by exploiting rationality.

Conclusion

Misinformation attacks can be computed in polynomial time by exploiting rationality.

New mechanisms are needed to combat misinformation attacks!