

Course:

Insights from Game Theory into Social Behavior

Instructors:

Moshe Hoffman, Erez Yoeli

TAs:

Zoe Hitzig, Maxim Massenkoff, Jason Nemirow

Let's discuss some puzzles of social behavior...

## Puzzle 1:

Norm against **Chemical Weapons**

See **The Oatmeal comic** about chemical weapons.

Why were chemical weapons the red line?

Why not 100,000 deaths?

Why not wanton murder of civilians?

This norm reared its head many times before...

“These flamethrowers were used to kill Japanese holed into pillboxes, buildings and caves. “  
-Wiki entry for Battle of Iwo Jima



Image is in the [public domain](#).

## Why flame throwers?

“A strong military case was made for the use of gas before America’s attack on the island of Iwo Jima; Japanese defenders in caves and tunnels would have been particularly vulnerable. Franklin Roosevelt rejected the idea.”

-The Economist, The History of Chemical Weapons



Image is in the [public domain](#).

>26,000 American casualties, Iwo Jima, WWII

“Because all the civilians had been evacuated,  
there were no civilian casualties at Iwo Jima”

## Video Clip: Japanese Soldier Burnt Alive by Flame Thrower

Note: This video contains content that is not suitable for all ages. You must be 18 years and over to view the content.

More humane than chemical weapon?

Thus,

- Chemical weapons would have saved lives!
- No civilian casualties either way!
- Not obviously “more humane”

In general:

- 1) Where do such inefficient norms come from?
- 2) What types of norms occur?

Puzzle 2:

Apologies

# Salala Pakistan, Nov 26 2011: US accidentally killed 24 Pakistani Soldiers



Image is in the public domain.

➔ Pakistan closed supply routes, until we apologized

Image removed to copyright restrictions.

[View](#) a map of NATO supply routes through Pakistan.

➔ ~\$1 billion in extra shipping fees



Image courtesy of [DayLove](#) on Flickr. CC BY-NC-SA

Until July 3 2012

“We are sorry for the losses suffered by the Pakistani military”

-Hillary Clinton

Immediately after...

“...the ground supply lines into Afghanistan are opening”

Why wouldn't US just say sorry?

-mere words?

-worth a billion?

Why would Pakistan care?

-wouldn't US "fake it"?

More generally:

- 1) **Why** do mere words matter?
- 2) **When** do mere words matter?
- 3) What about other symbolic actions (e.g. **coronations, graduations, handshakes**, etc)?

Puzzle 3:

Why do we consider transgressions of **commission** worse than those of **ommission**?



Notice:

- Batman's intention is the same
- The outcome is the same
- But Batman (and presumably the viewer) thinks omission less bad

- 1) **Why** is omission viewed differently from commission?
- 2) Is this distinction something we should **legally** respect or overcome?

More generally:

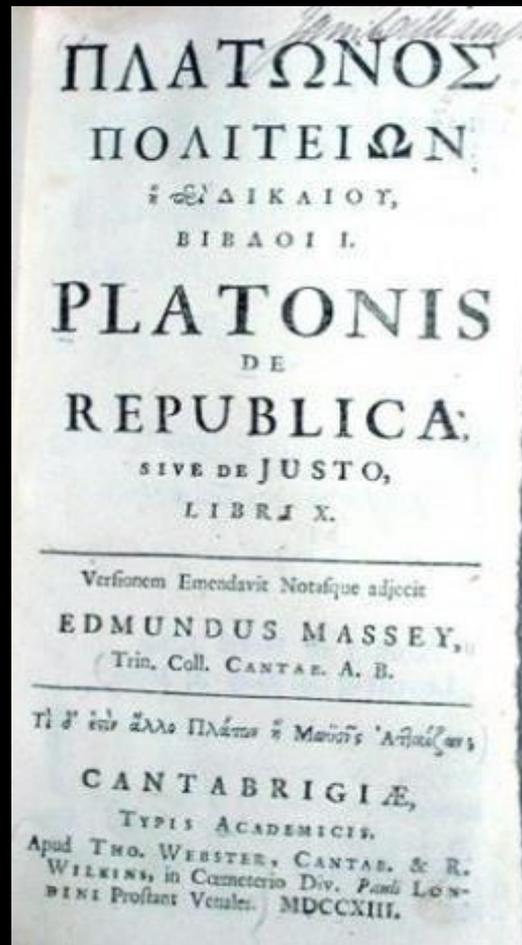
- 1) **Where** do our moral intuitions come from?
- 2) Do they make a good basis for **law**?

Puzzle 4:

Where do “rights” come from?



# Might?



# The “state of nature”?

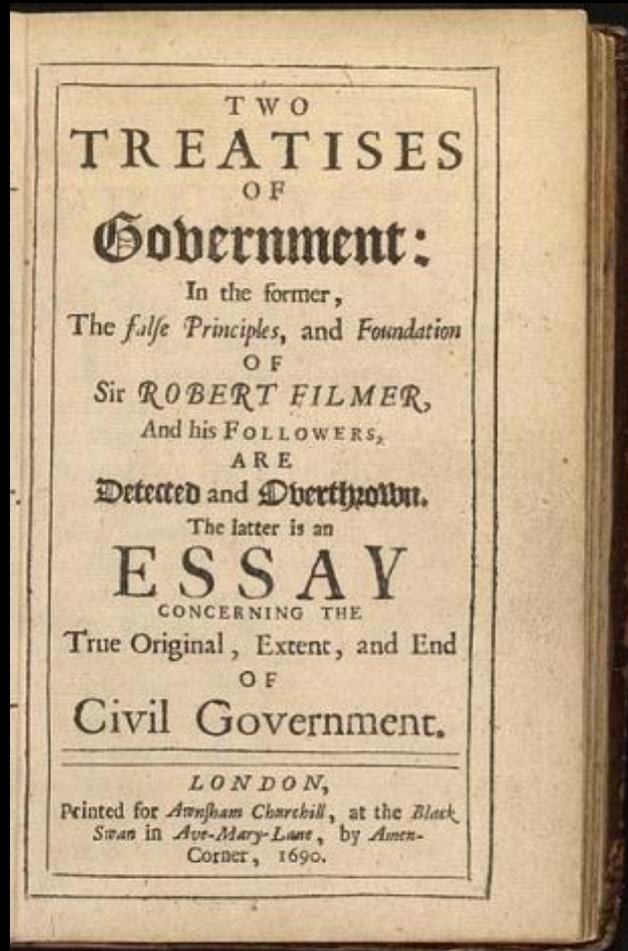
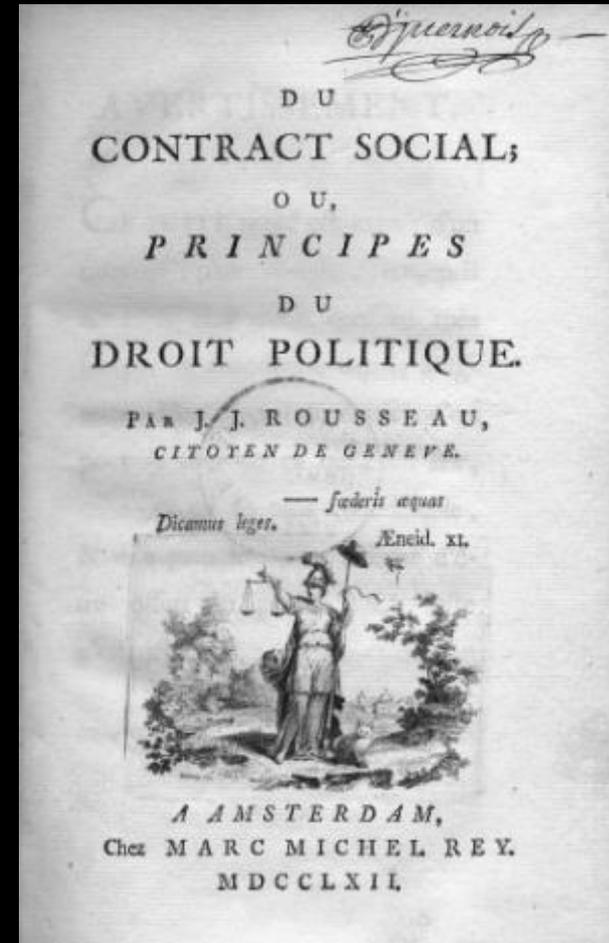
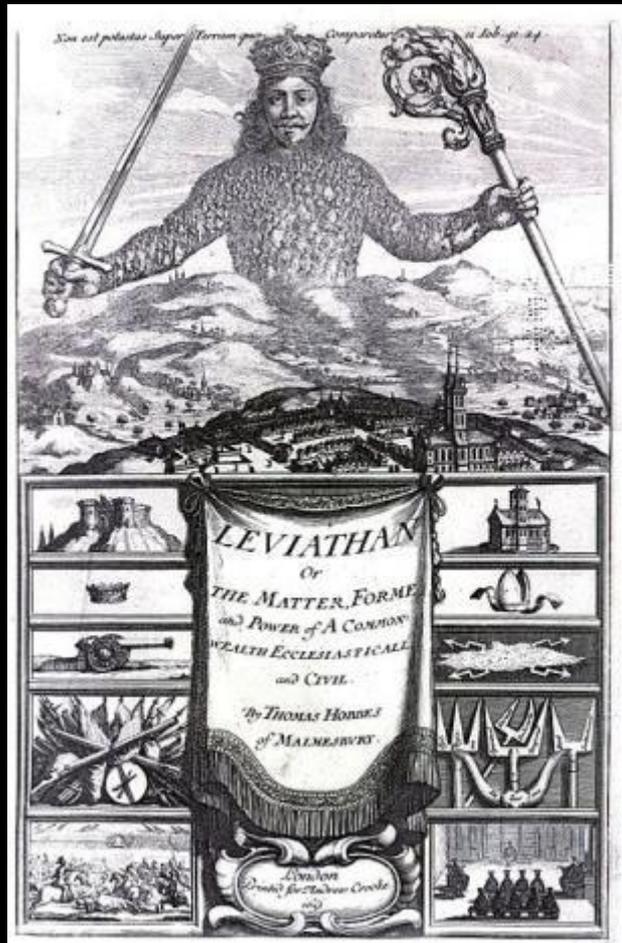


Image is in the public domain.

# A “social contract”?



What does this mean?

Where **DO** rights come from?

Puzzle 5:

Why do we speak **indirectly**?



- 1) Did this line introduce any **doubt** as to whether a bribe was offered?
- 2) Would a **good cop** be any less likely than a corrupt cop to “get it”?
- 3) Was the goal to prevent “**proof**” that a bribe was offered?

More generally,

- 1) Why do we communicate in this **inefficient** way?
- 2) When is it **important** to do so?

In this class...

We will explain such puzzles using game theory

In this class...

We will explain such puzzles using game theory

What is game theory?

The simplest “game” can be represented by the following “payoff matrix”

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

Player 1 chooses between two actions

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

Player 2 **simultaneously** chooses between 2 actions

L

R

U

5, 6

8, 4

D

3, 2

0, -3

The **payoffs** to player 1 are determined by her action as well as the action of player 2

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

The payoffs to player 2 are determined by her action as well as the action of player 1

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

This game can be “solved” by finding the “Nash equilibria”

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

(U, L) is a Nash Equilibrium b/c neither can benefit by **unilaterally deviating**

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

“Prediction” of game theory:

If both “expected” (U,L), both would play (U,L)!

(Nash is “self enforcing”)

(U,R) is NOT a Nash Equilibrium b/c 2 can benefit by unilaterally deviating to L

	L	R
U	5, 6	8, 4
D	3, 2	0, -3

Game theory “predicts”:

If both expected (U,R), player 2 would deviate!

(I.e. if not Nash, cannot be “stable”)

Nash makes sense (*arguably*) if...

-Uber-rational

-Calculating

# Such as Auctions...



# Or Oligopolies...



Image courtesy of [afagen](#) on Flickr. CC BY NC-SA



Image courtesy of [longislandwins](#) on Flickr. CC-BY

But why would game theory matter for our puzzles?

Norms/rights/morality are not **chosen**; rather...

We **believe** we have rights!

We **feel** batman would be worse if he killed

We use an innuendo because it **feels** awkward to *explicitly* say inappropriate.

Apologies matter because recipients **feel** nice when they hear them

But...

From where do these feelings/beliefs come?

Our thesis (in a few steps):

Thesis:

Feelings/beliefs that “do better” become “more frequent”

Thesis:

Feelings/beliefs that “do better” become “more frequent”

Because of “evolution”

(people with certain beliefs or preferences die out?)

Thesis:

Feelings/beliefs that “do better” become “more frequent”

Because of “selective imitation”

(people with certain beliefs or preferences are more likely to be imitated?)

Thesis:

Feelings/beliefs that “do better” become “more frequent”

Because of “reinforcement learning”

(certain beliefs or preferences are held onto more tenaciously?)

Thesis:

Feelings/beliefs that do better become more frequent

→ behavior ends up consistent with Nash

## Thesis:

Feelings/beliefs that do better become more frequent

→ behavior ends up consistent with Nash  
(o.w. feelings/beliefs would **change!**)

## Thesis:

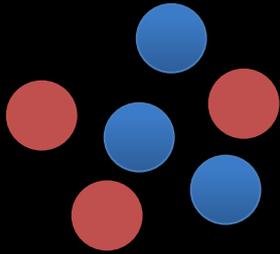
Feelings/beliefs that do better become more frequent

→ behavior ends up consistent with Nash  
(o.w. feelings/beliefs would change!)

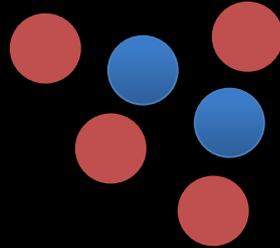
(even though not **aware** of game, or of where feelings/beliefs come from)

In detail...

What is the **key assumption** in evolution, imitation, reinforcement learning?



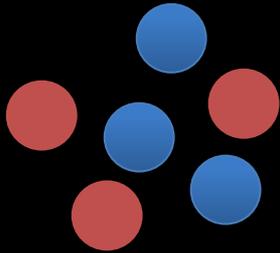
T=0



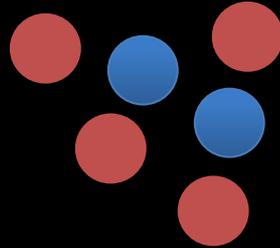
T=1

More successful **traits** reproduce faster

Evolution



T=0



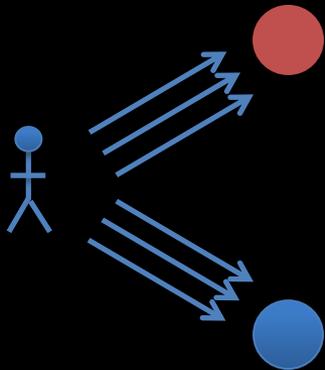
T=1

More successful **traits** more likely to be imitated

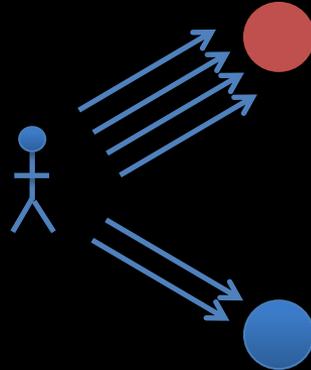
Selective  
Imitation

## Reinforcement Learning

T=0

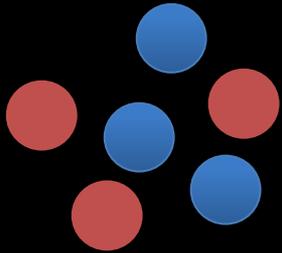


T=1

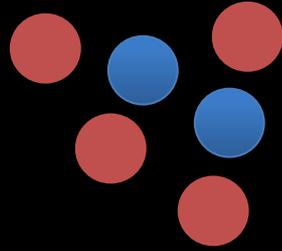


More successful **behaviors**  
held more tenaciously

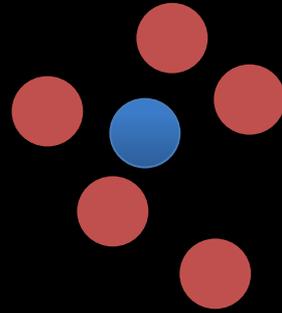
All 3 processes → “optimal behaviors”!



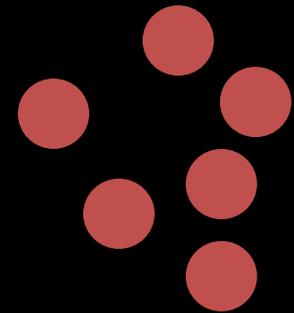
T=0



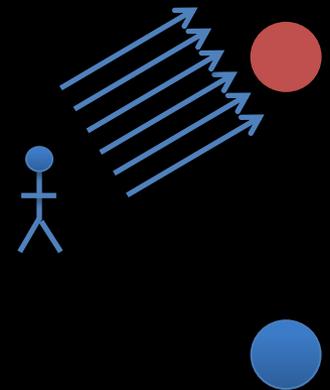
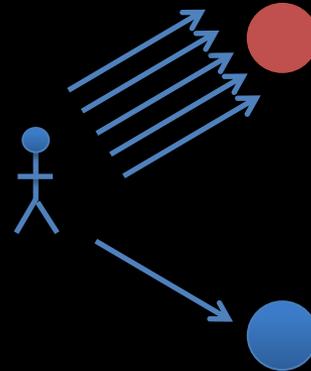
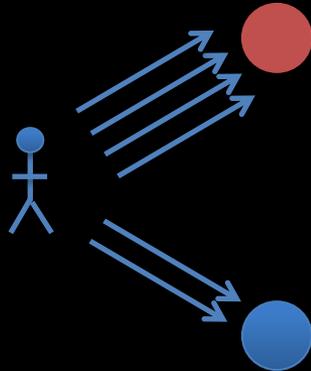
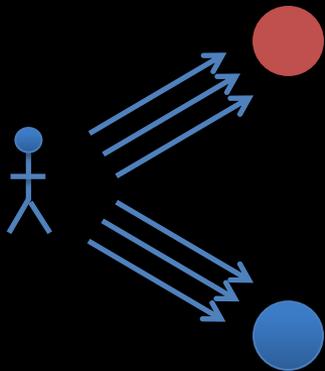
T=1



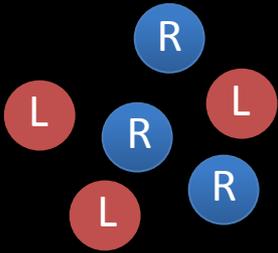
T=2



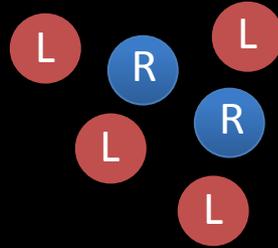
T=3



Also true in a game...

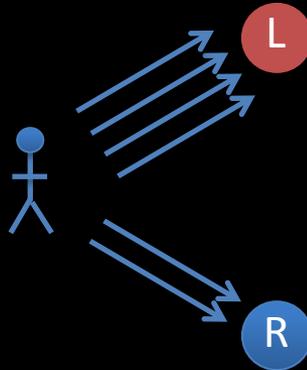
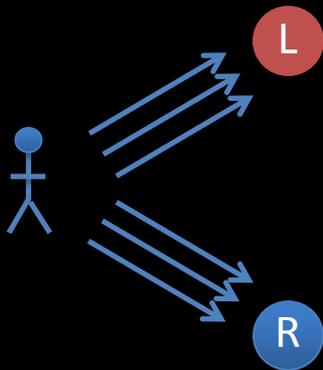


T=0

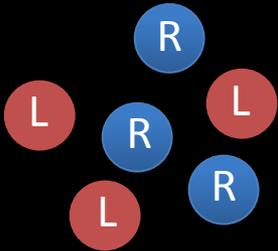


T=1

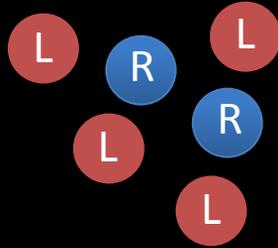
More successful **strategies** become more frequent



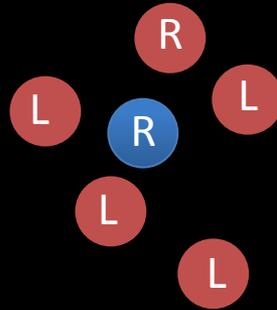
And eventually take over...



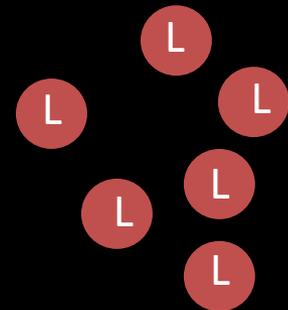
T=0



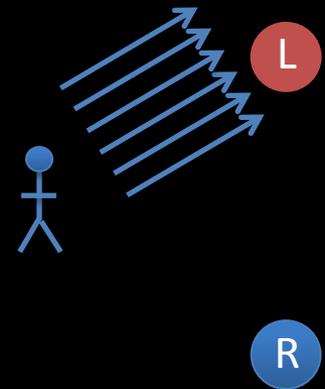
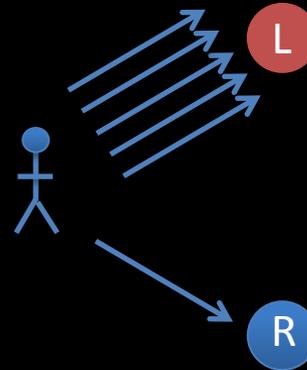
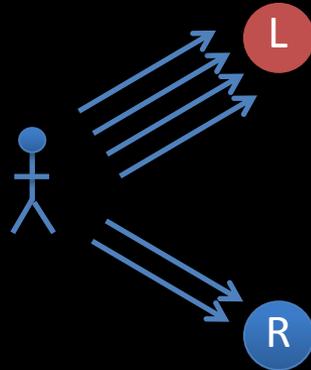
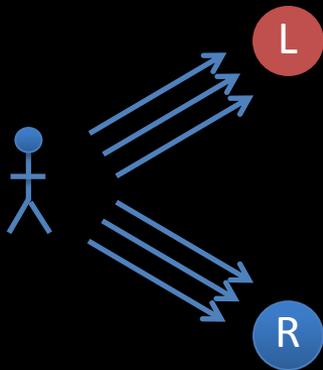
T=1



T=2

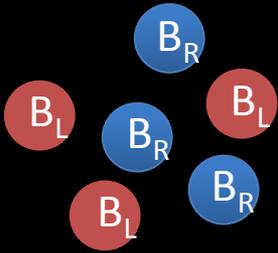


T=3

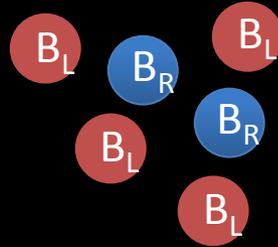


What if **beliefs/feelings** are being  
learned/evolving instead of “**strategies**”?

Suppose  $B_L$  is belief that causes action L to be taken

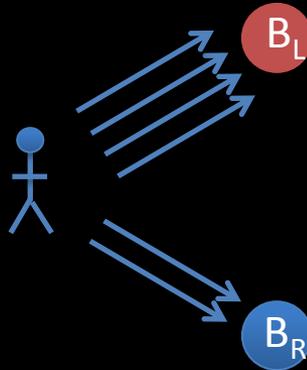
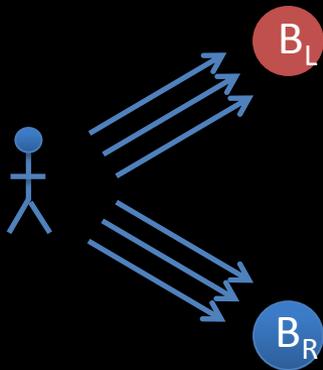


T=0

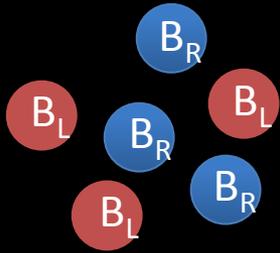


T=1

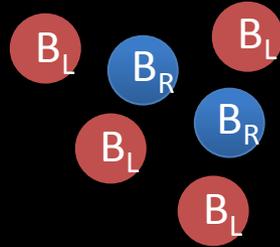
Feelings/beliefs that do better become more frequent



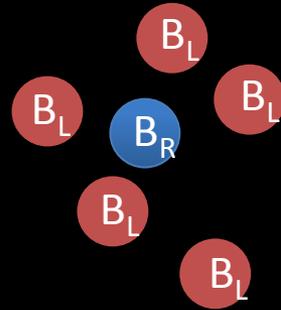
# Behavior ends up consistent with Nash



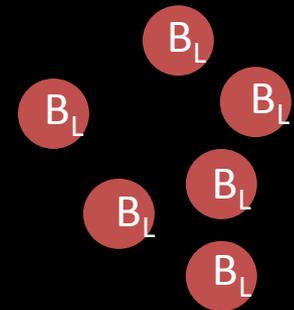
T=0



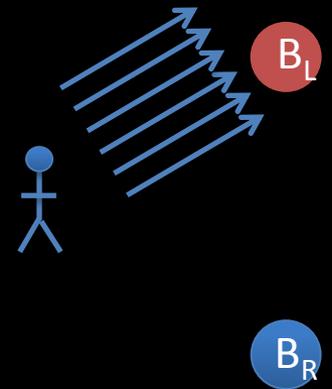
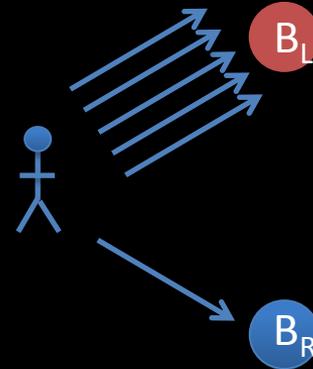
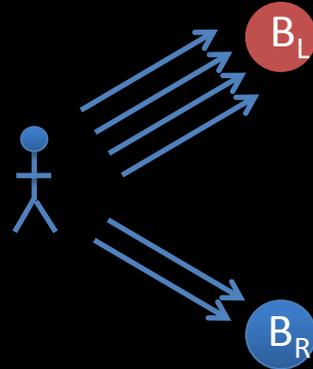
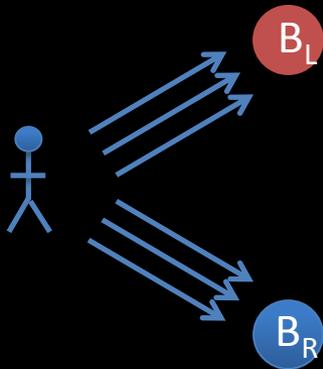
T=1



T=2



T=3



(even though not **conscious** of Nash, or of where feelings come from)

MIT OpenCourseWare  
<http://ocw.mit.edu>

14.11 Insights from Game Theory into Social Behavior  
Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.