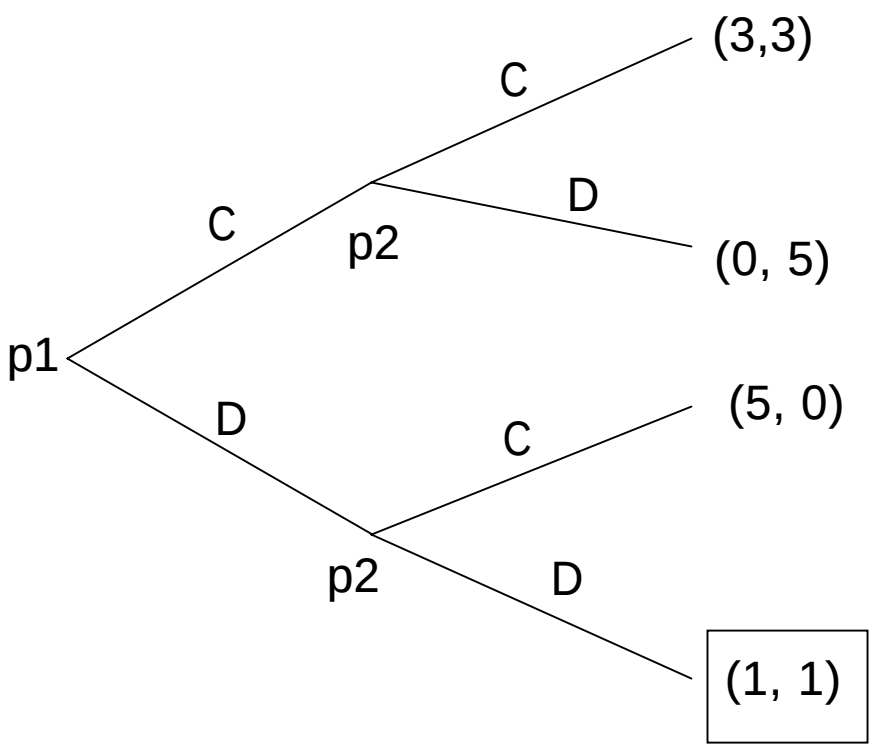
Iterated Prisoner's Dilemma

CS 790R — Spring 2005

presented by Jeff Wallace

Extended Form



- imperfect information
- Nash equilibrium
- backward induction finds (1,1)

Single Iteration

p2 / p1	C	D
C	(3,3)	(5,0)*
D	*(5,0)	*(1,1)*

- matrix view
- hold one player's move constant and mark choice with *
- one perfect NE

One Iteration – General Form

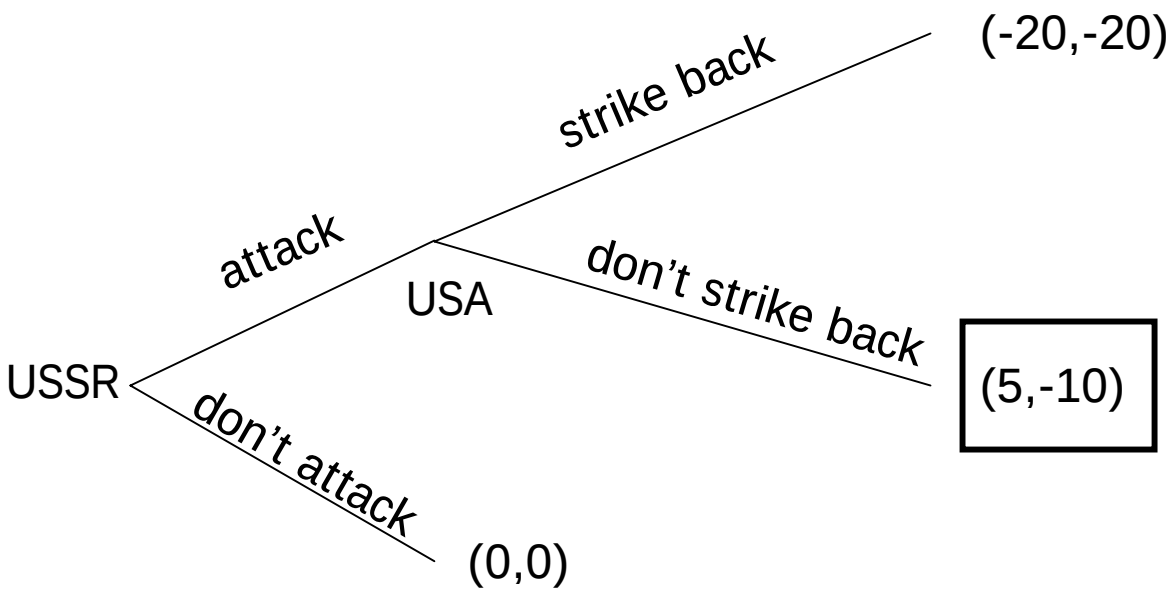
p1 \ p2	C	D
C	(R,R)	(S,T)*
D	*(T,S)	*(P,P)*

- dilemma where T>R>P>S and 2R > T+S
- (R,S,T,P) = (3,0,5,1) or $(1,0,{}^{5}/_{3},{}^{1}/_{3})$
- aka (p,q) = $({}^{5}/_{3},{}^{1}/_{3})$

Requirements for NE

- Reinforcement of beliefs: players believe they know how other will play, which is self-enforcing
- No incentive for “do over.” Players are rational.

Payoffs Dictate Equilibriums



- nuclear holocaust "game" of the cold war
- why didn't the perfect NE occur?

Iterated Prisoner's Dilemma

- Extended form (view from Elmo device)
- Backwards induction shows one perfect NE resulting from 'DD', with average payoff = (1,1)
- Proof: if either player deviates, he earns zero at each step, therefore there's no incentive to deviate?

Iterated PD (CONT'D)

What about...

- Play 'C' in period 1, then keep playing 'C' in subsequent periods until p2 plays 'D'. Thereafter, play 'D'. Same for p2.
  - this gives outcome C-C, C-C,...,C-C with average payoff = (3,3). It's another NE!
  - proof: can either player deviate? A: no, because if p1 plays 'D', p2 will switch to 'D' forever, leading to long-term avg payoff = (1,1). P2 can't deviate for same reason.

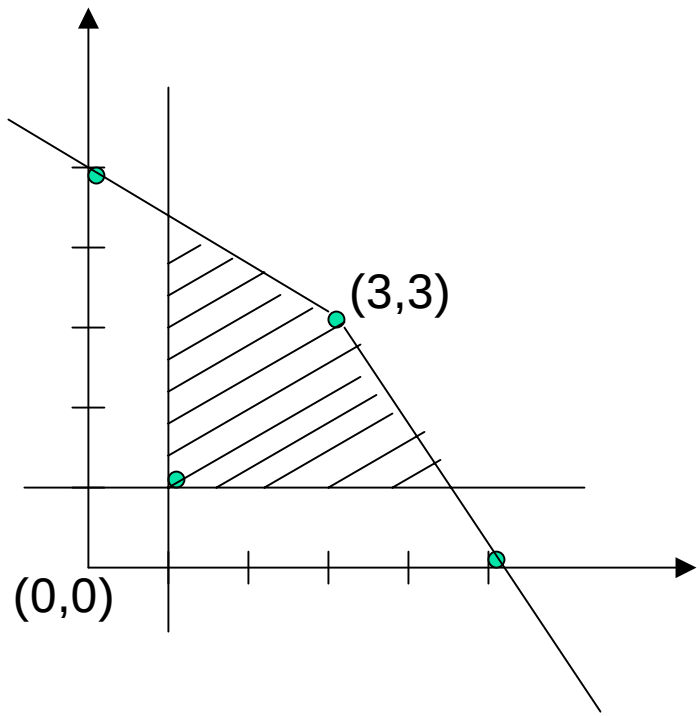
Iterated PD (CONT'D)

What about...

- Play C,D,C,D...as long as opponent plays same. If opponent deviates, play 'D' forever.
  - this gives outcome C-C, D-D,... with average payoff = (2,2). It's another NE!
  - proof: can either player deviate? A: no. Same reasons as before.

Individually Rational Zone

- Convex hull within which players get at least (1,1).



- Infinite number of NEs within IRZ
- this leads to mixed strategies as NE (demonstrate mixed strat on board)

TFT is a Mixed Strategy

- Ex. C-C, C-C, D-C, C-D, C-C,...
  - Result = (3,3)+(3,3)+(5,0)+(0,5)+(3,3) =
  - average payoff = $(2\ ^4/_5,\ 2\ ^4/_5)$
  - This is within IRZ (convex hull)
- PAV strategy is similar. It also lies within IRZ

Other Mixed Strategy Results

- ALL-D approaches (1,1) and is therefore eventually self-destructive. It almost gets wiped out. “Good parasite, but needs a healthy host.”
- RAND is overcome by “nice” strategies.
- TFT dominates in environments without noise, but PAV dominates with noise.
- ALL-C marginally survives without noise, but becomes extinct with noise.