Richer Representations: Beyond Normal and Extensive Forms

Richer Game Representations

- Normal, extensive forms not always suitable for modeling large or realistic game-theoretic settings
- Infinite games
  - Repeat finite game infinitely (such as Prisoner's Dilemma)
  - Infinite action space
- Even if finite, games can quickly become too large to reason about with NF, EF

Richer Game Representations

- Luckily, not usually interested in arbitrary strategic settings
- Highly structured situations
  - Repeated play of small games (i.e. game unfolds over time)
  - Nature of problem domain (e.g. number of agents interacting at one time is small)
- In this module, will look at repeated games. Next module: stochastic, bayesian games

Finitely Repeated Games

Twice-played Prisoner's

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

$\Rightarrow$

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Figure 6.1: Twice-played Prisoner's Dilemma.

Twice-played Prisoner's

- One way to disambiguate: represent in extensive form

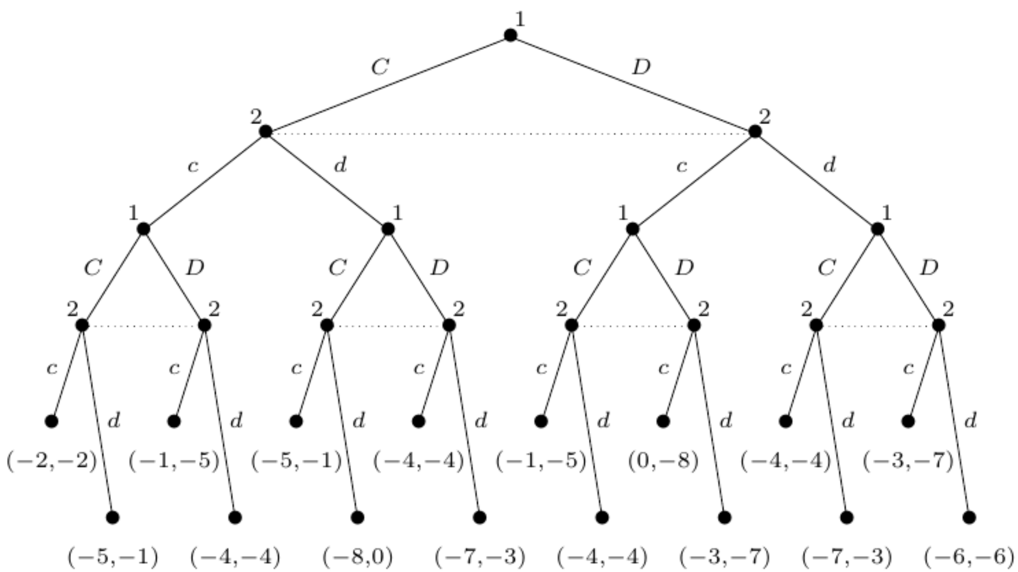

Figure 6.2: Twice-played Prisoner's Dilemma in extensive form.

Twice-played Prisoner's

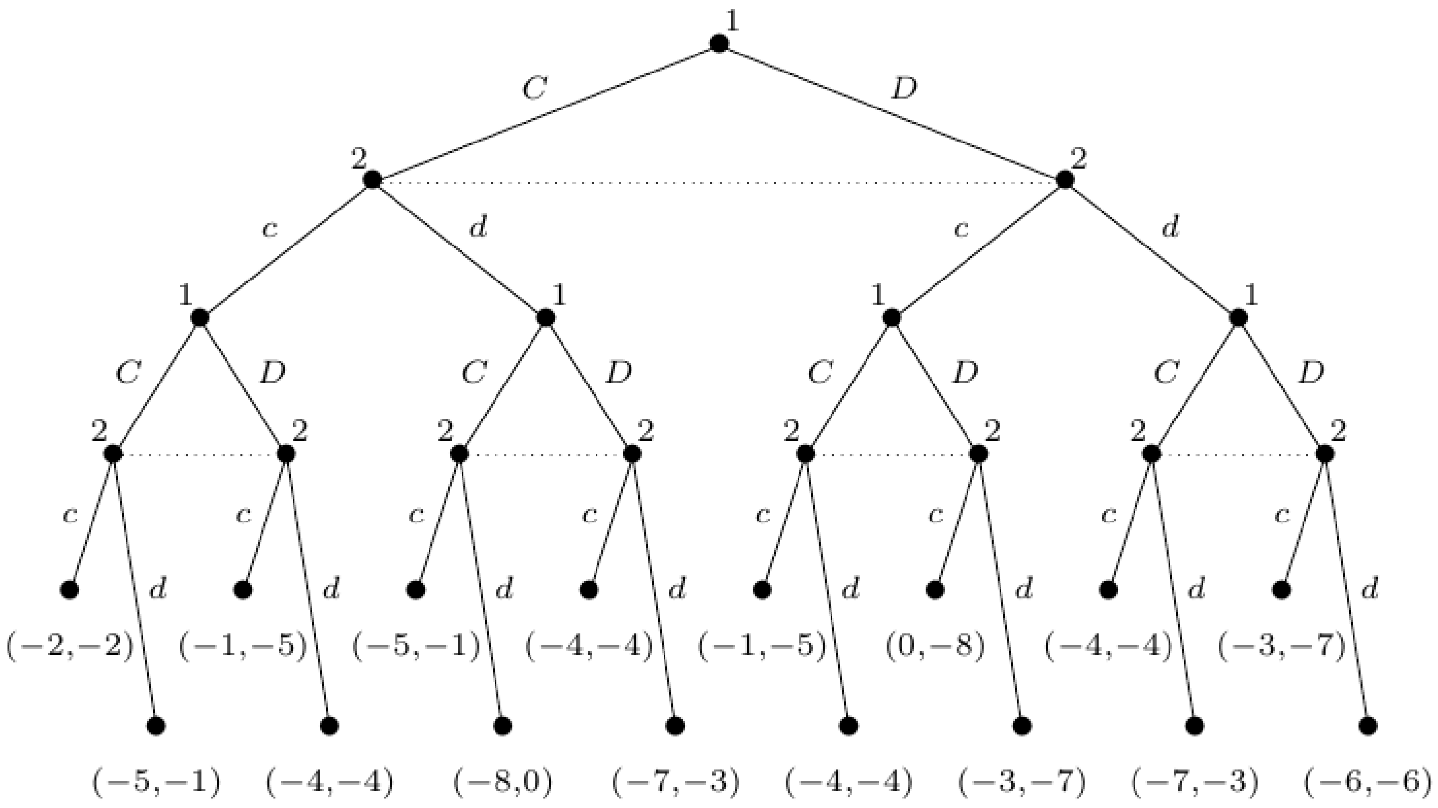

Figure 6.2: Twice-played Prisoner's Dilemma in extensive form.

Repeated Games

- Strategy space in repeated game is much richer
- *Stationary strategy*
- Action or mixture can depend on the history thus far
- Recall backward induction on the centipede game
- Similar argument says one should always defect in each round: but empirically and theoretically, there are problems with this

Infinitely Repeated Games

Infinite Repetition

- Repeat normal-form game for an infinite number of repetitions
- No longer can transform into extensive form, would result in infinite tree
- How to define payoffs?

Two Possibilities

Definition 6.1.1 (Average reward) *Given an infinite sequence of payoffs* $r_i^{(1)}, r_i^{(2)}, \ldots$ *for player* i, *the* average reward *of* i *is*

$$\lim_{k \to \infty} \frac{\sum_{j=1}^{k} r_i^{(j)}}{k}.$$

Definition 6.1.2 (Discounted reward) *Given an infinite sequence of payoffs* $r_i^{(1)}, r_i^{(2)}, \ldots$ *for player* i, *and a discount factor* β *with* $0 \leq \beta \leq 1$, *the* future discounted reward *of* i *is* $\sum_{j=1}^{\infty} \beta^j r_i^{(j)}$.

Infinitely Repeated Prisoners' Dilemma

- *Tit-for-tat (TfT):* Start by cooperating; thereafter, pick in round $j + 1$ the action chosen by the other player in round j
  - Hard to beat this strategy; won several competitions
  - If discount factor is large enough, is a Nash equilibrium
- *Trigger* strategy: Start by cooperating; if other player ever defects, defect forever
  - Also Nash equilibrium for large enough discount

The Folk Theorem

- Can we characterize the Nash equilibria?
- What rewards are possible in a Nash equilibrium?
- Average rewards attainable in equilibrium are those available in mixed strategies in original game, as long as at least the minmax value
- The idea is: Players can respond to a deviation by punishing infinitely (adopting minmax strategy against the deviant player)

The Folk Theorem

- Game *G=(N,A,u);* payoff profile $r = (r_1, ..., r_n)$; and $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$.

Definition 6.1.3 (Enforceable) *A payoff profile* $r = (r_1, r_2, \ldots, r_n)$ *is* enforceable *if* $\forall i \in N$, $r_i \geq v_i$.

Definition 6.1.4 (Feasible) *A payoff profile* $r = (r_1, r_2, \ldots, r_n)$ *is* feasible *if there exist rational, nonnegative values* α_a *such that for all* i*, we can express* r_i *as* $\sum_{a \in A} \alpha_a u_i(a)$*, with* $\sum_{a \in A} \alpha_a = 1$.

The Folk Theorem

Theorem 6.1.5 (Folk Theorem) *Consider any n-player normal-form game G and any payoff profile $r = (r_1, r_2, \ldots, r_n)$.*

1. *If r is the payoff profile for any Nash equilibrium s of the infinitely repeated G with average rewards, then for each player i, r_i is enforceable.*
2. *If r is both feasible and enforceable, then r is the payoff profile for some Nash equilibrium of the infinitely repeated G with average rewards.*

Bounded Rationality

Recall Repeated Prisoners'

	C	D
C	$3, 3$	$0, 4$
D	$4, 0$	$1, 1$

Figure 6.3: Prisoner's Dilemma game.

Recall Repeated Prisoners'

- What could explain that humans often cooperate (early on) in repeated Prisoners' Dilemma game?
- One idea: players are not entirely rational
- "bounded rationality" -- various models exist
  - approximate Nash equilibria
  - Restriction of players' strategy space to automata of various types

Finite-State Automata

Definition 6.1.6 (Automaton) *Given a game $G = (N, A, u)$ that will be played repeatedly, an automaton M_i for player i is a four-tuple $(Q_i, q_i^0, \delta_i, f_i)$, where:*

- *Q_i is a set of states;*
- *q_i^0 is the start state;*
- *$\delta_i : Q_i \times A \mapsto Q_i$ is a transition function mapping the current state and an action profile to a new state; and*
- *$f_i : Q_i \mapsto A_i$ is a strategy function associating with every state an action for player i.*

$$a_i^t = f_i(q_i^t)$$
$$q_i^{t+1} = \delta_i(q_i^t, a_1^t, \ldots, a_n^t)$$

Examples

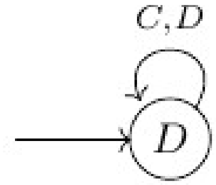

Figure 6.4: An automaton representing the repeated *Defect* action.

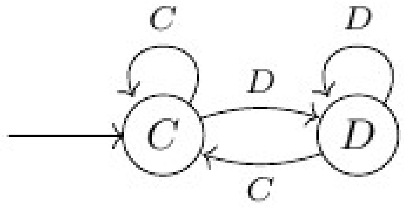

Figure 6.5: An automaton representing the *Tit-for-Tat* strategy.

Machine Game

Definition 6.1.7 (Machine game) *A two-player machine game* $G^M = (\{1,2\}, \mathcal{M}, G)$ *of the* k*-period repeated game* G *is defined by:*

- *a pair of players* $\{1,2\}$*;*
- $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$*, where* $\mathcal{M}_i$ *is a set of available automata for player* i*; and*
- *a normal-form game* $G = (\{1,2\}, A, u)$.

A pair $M_1 \in \mathcal{M}_1$ *and* $M_2 \in \mathcal{M}_2$ *deterministically yield an outcome* $o^t(M_1, M_2)$ *at each iteration* t *of the repeated game. Thus,* G^M *induces a normal-form game* $(\{1,2\}, \mathcal{M}, U)$*, in which each player* i *chooses an automaton* $M_i \in \mathcal{M}_i$*, and obtains utility* $U_i(M_1, M_2) = \sum_{t=1}^{k} u_i(o^t(M_1, M_2))$.

Bounded Rationality: Limiting States

- Idea: Automata with fewer states represent *simpler* strategies
  - Can bound rationality by bounding the number of states in the automaton
- With *k*-repeated Prisoners', if the automata are restricted to less than *k* states, the constant-defect strategy does not yield a symmetric equilibrium. But Tit-for-Tat does

Theorem 6.1.8 *For any integer x, there exists an integer k_0 such that for all $k > k_0$, any machine game $G^M = (\{1,2\}, \mathcal{M}, G)$ of the k-period repeated Prisoner's Dilemma game G, in which $k^{1/x} \leq \min\{S(\mathcal{M}_1), S(\mathcal{M}_2)\} \leq \max\{S(\mathcal{M}_1), S(\mathcal{M}_2)\} \leq k^x$ holds has a Nash equilibrium in which the average payoffs to each player are at least $3 - \frac{1}{x}$.*

Bounded Rationality: Cost of Complexity

- Idea: Give players a disutility for complexity

Definition 6.1.10 (Lexicographic disutility for complexity) *Agents have* lexicographic disutility for complexity *in a machine game if their utility functions* $U_i(\cdot)$ *in the induced normal-form game are replaced by preference orderings* $\succeq_i$ *such that* $(M_1, M_2) \succ_i (M_1', M_2')$ *whenever either* $U_i(M_1, M_2) > U_i(M_1', M_2')$ *or* $U_i(M_1, M_2) = U_i(M_1', M_2')$ *and* $s(M_i) < s(M_i')$.

Example

- Infinitely repeated Prisoners'
- Player 2 using trigger strategy
- Player 1 cannot achieve higher payoff other than playing trigger herself
- But Player 1 can achieve same utility by always cooperating.
- So (Trigger, Trigger) is not a Nash equilibrium

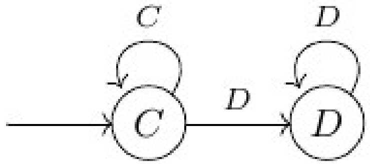

Figure 6.6: An automaton representing the Trigger strategy.

Other Considerations

- Computing best-response automata
  - Problem of verifying best-response automaton is NP-complete

Theorem 6.1.13 *Given a machine game $G^M = (\{1, 2\}, \mathcal{M}, G)$ of the limit average infinitely repeated Prisoner's Dilemma game G, an automaton M_2, and an integer k, the problem of computing a best-response automaton M_1 for player 1, such that $s(M_1) \leq k$, is NP-complete.*

- Finite automata to Turing machines
  - Best response may not even be a Turing machine computable strategy