Behavioral Game Theory

Modeling Strategic Behavior as a Machine Learning Problem

Kevin Leyton-Brown

University of British Columbia
Canada CIFAR AI Chair, Amii

James R. Wright

University of Alberta
Canada CIFAR AI Chair, Amii

Lecture Overview

Motivation

Behavioral Game Theory

Normative vs. Descriptive

The solution concepts we have studied so far have been **normative**: They identify outcomes that satisfy some assumptions or criteria:

- Pareto optimality
  - No agent can improve their utility without reducing someone else's
- Nash equilibrium
  - Accurate beliefs about accurate beliefs about...("rational expectations")
  - Mutual best response ("rational response")

Normative vs. Descriptive

The solution concepts we have studied so far have been **normative**: They identify outcomes that satisfy some assumptions or criteria:

- Pareto optimality
  - No agent can improve their utility without reducing someone else's
- Nash equilibrium
  - Accurate beliefs about accurate beliefs about...("rational expectations")
  - Mutual best response ("rational response")

But often, we want to answer the **descriptive** question: What will *actually happen?*

Game Theory is Not Descriptive

- Nash equilibrium is often treated as a descriptive prediction
- But its predictions are often pretty counter-intuitive
  - That's because they don't actually do a great job of predicting real behavior

Game Theory is Not Descriptive

- Nash equilibrium is often treated as a descriptive prediction
- But its predictions are often pretty counter-intuitive
  - That's because they don't actually do a great job of predicting real behavior

Examples: [Goeree & Holt 2001]

1. Traveller's Dilemma
   - Essentially nobody plays the unique equilibrium of 2
   - If you make the penalty large, people play much closer to Nash equilibrium
   - But the size of the penalty does not affect the equilibrium

Game Theory is Not Descriptive

- Nash equilibrium is often treated as a descriptive prediction
- But its predictions are often pretty counter-intuitive
  - That's because they don't actually do a great job of predicting real behavior

Examples: [Goeree & Holt 2001]

1. Traveller's Dilemma
   - Essentially nobody plays the unique equilibrium of 2
   - If you make the penalty large, people play much closer to Nash equilibrium
   - But the size of the penalty does not affect the equilibrium
2. Asymmetric Matching Pennies
   - Increasing the payoff for one of a single player's action doesn't change their own unique equilibrium strategy
   - But it frequently changes that player's behavior!

Lecture Overview

Motivation

Behavioral Game Theory

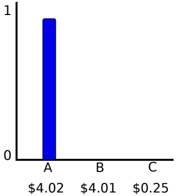
Behavioral Game Theory

- **Behavioral game theory** aims to solve the **descriptive** problem
  - Proposes models to better explain or predict human behavior
- Often by relaxing assumptions, e.g.:
  - Rational expectations
  - Rational response
  - Dynamic consistency
  - ...

Behavioral Game Theory

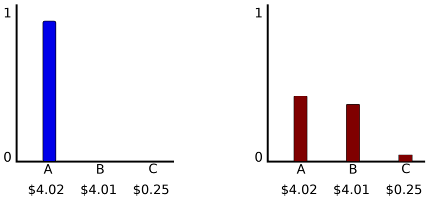
- **Behavioral game theory** aims to solve the **descriptive** problem
  - Proposes models to better explain or predict human behavior
- Often by relaxing assumptions, e.g.:
  - **Rational expectations**
  - **Rational response**
  - Dynamic consistency
  - ...

1. Relaxing Best Response



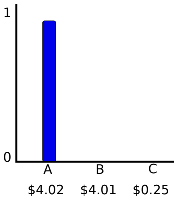
- Best response: Maximum utility action is **always** played

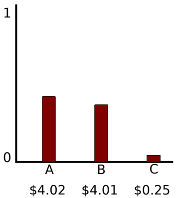
1. Relaxing Best Response



- Best response: Maximum utility action is **always** played
- Quantal response: High-utility actions played often than low-utility actions

1. Relaxing Best Response



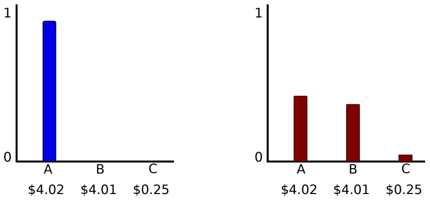


- Best response: Maximum utility action is **always** played
- Quantal response: High-utility actions played often than low-utility actions

Usual specification: Logistic best response ("softmax")

$$QBR_i(s_{-i};\lambda)(a_i) = \frac{\exp(\lambda u_i(a_i, s_{-i}))}{\sum_{a_i' \in A_i} \exp(\lambda u_i(a_i', s_{-i}))}$$

1. Relaxing Best Response



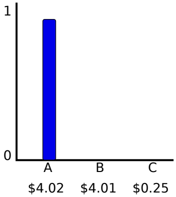
- Best response: Maximum utility action is **always** played
- Quantal response: High-utility actions played often than low-utility actions

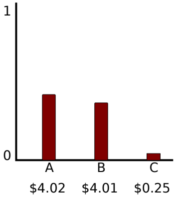
Usual specification: Logistic best response ("softmax")

$$QBR_i(s_{-i};\lambda)(a_i) = \frac{\exp(\lambda u_i(a_i, s_{-i}))}{\sum_{a'_i \in A_i} \exp(\lambda u_i(a'_i, s_{-i}))}$$

λ represents sensitivity to differences in utility.

Interpreting Quantal Response

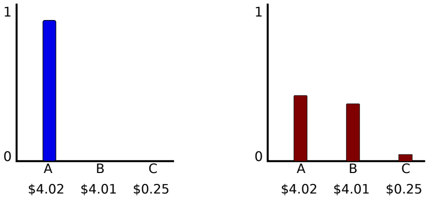




Interpretations of quantal response:

1. People choose randomly, but are more likely to choose high-utility actions

Interpreting Quantal Response



Interpretations of quantal response:

1. People choose randomly, but are more likely to choose high-utility actions
2. People maximize their utility, but their utility is randomly "shocked":
   - We observe $v(a_i)$ for each action a_i
   - Agents choose $\arg\max_i v(a_i) + \xi_\mathbf{i}$
   - where each ξ_i is a random variable

Model: Quantal Response Equilibrium

Definition (Quantal Response Equilibrium)

A strategy profile s is a **quantal response equilibrium** (QRE) with precisions λ if every agent is simultaneously quantally responding to the profile of the other agents' strategies, i.e.

$$\forall i \in N : s_i = QBR_i(s_{-i}; \lambda)$$

Model: Quantal Response Equilibrium

Definition (Quantal Response Equilibrium)

A strategy profile s is a **quantal response equilibrium** (QRE) with precisions λ if every agent is simultaneously quantally responding to the profile of the other agents' strategies, i.e.

$$\forall i \in N : s_i = QBR_i(s_{-i}; \lambda)$$

- Note that agents still have rational expectations: they are responding to the **real strategies** of the other agents.
- Equilibrium selection is still a question:
  - There can be multiple QREs for a given precision
  - It's not clear how agents would arrive at a QRE

2. Relaxing Rational Expectations

Can instead relax rational expectations:

- Doesn’t seem plausible that people would all know each other’s actual strategies
- Especially doesn’t seem plausible that agents would have accurate high-order beliefs to unlimited levels of recursion!
- But then what should we assume that people believe?

Iterative Strategic Reasoning

Every agent performs some **finite number** of steps of strategic reasoning:

- level-0: Some default, nonstrategic distribution of play (often uniform)
- level-1: Best response to level-0 players
- level-2: Best response to level-1, or to level-1 and level-0

$$\vdots$$

- Level-k: Best response to level $k-1$, or to levels $\{0, 1, \ldots, k-1\}$

Model: Level-k

Definition (Level-k)

A strategy profile s is the prediction of a **level-k** model with parameters $\alpha_0, \alpha_1, \ldots, \alpha_K$ and level-0 strategies $\pi^{\text{Lk}}_{i,0}$ if

$$\forall i \in N : s_i = \sum_{k=0}^{K} \alpha_k \pi^{\text{Lk}}_{i,k}$$

where $\pi^{\text{Lk}}_{i,k} = BR_i(\pi^{\text{Lk}}_{-i,k-1})$ for all $k > 0$.

- Every agent has a fixed “level” representing the number of steps of strategic reasoning they can perform
  - They assume that every other agent performs exactly one step fewer
- Parameter $\pi^{\text{Lk}}_{i,0}$ represents the level-0 strategy for each agent
- Parameters $\alpha_0, \ldots, \alpha_K$ represent the frequency of levels

Model: Cognitive Hierarchy

A common variation:

- Agents know that other agents might perform any number of steps of reasoning (less than their own)
- Agents know the relative proportions of all lower levels

Model: Cognitive Hierarchy

A common variation:

- Agents know that other agents might perform any number of steps of reasoning (less than their own)
- Agents know the relative proportions of all lower levels
- **Question:** Why would we make this assumption?

Model: Cognitive Hierarchy

A common variation:

- Agents know that other agents might perform any number of steps of reasoning (less than their own)
- Agents know the relative proportions of all lower levels
- **Question:** Why would we make this assumption?

Definition (Cognitive Hierarchy)

A strategy profile s is the prediction of a **cognitive hierarchy** model with parameters $\alpha_0, \alpha_1, \ldots, \alpha_K$ and level-0 strategies $\pi_{i,0}^{\text{CH}}$ if

$$s_i = \sum_{k=0}^{K} \alpha_k \pi_{i,k}^{\text{CH}}$$

where $\pi_{i,k}^{\text{CH}} = BR_i(\pi_{-i,0:k-1}^{\text{CH}})$ for all $k > 0$, and $\pi_{i,0:k}^{\text{CH}} = \frac{\sum_{j=0}^{k} \alpha_j \pi_{i,j}^{\text{CH}}}{\sum_{j=0}^{k} \alpha_j}$.

Model: Quantal Cognitive Hierarchy (QCH)

Of course, you can easily relax both rational expectations and rational response:

Definition (Quantal Cognitive Hierarchy)

A strategy profile s is the prediction of a **cognitive hierarchy** model with parameters $\lambda, \alpha_0, \alpha_1, \ldots, \alpha_K$ and level-0 strategies $\pi_{i,0}^{\text{QCH}}$ if

$$s_i = \sum_{k=0}^{K} \alpha_k \pi_{i,k}^{\text{QCH}}$$

where $\pi_{i,k}^{\text{QCH}} = QBR_i(\pi_{-i,0:k-1}^{\text{QCH}}; \lambda)$ for all $k > 0$, and $\pi_{i,0:k}^{\text{QCH}} = \frac{\sum_{j=0}^{k} \alpha_j \pi_{i,j}^{\text{QCH}}}{\sum_{j=0}^{k} \alpha_j}$.

2a. Relaxing Uniform Level-0

Bonus assumption!

- Most existing work that uses level-k/cognitive hierarchy style models assumes that $\pi_{i,0}$ is the **uniform distribution**
- Others hand-pick a game-specific "default strategy" (e.g., 300 in Traveller's Dilemma)

2a. Relaxing Uniform Level-0

Bonus assumption!

- Most existing work that uses level-k/cognitive hierarchy style models assumes that $\pi_{i,0}$ is the **uniform distribution**
- Others hand-pick a game-specific “default strategy” (e.g., 300 in Traveller’s Dilemma)
- But uniform distribution is pretty implausible
- And hand-picked defaults are hard to justify, don’t scale to arbitrary games

2a. Relaxing Uniform Level-0

Bonus assumption!

- Most existing work that uses level-k/cognitive hierarchy style models assumes that $\pi_{i,0}$ is the **uniform distribution**
- Others hand-pick a game-specific "default strategy" (e.g., 300 in Traveller's Dilemma)
- But uniform distribution is pretty implausible
- And hand-picked defaults are hard to justify, don't scale to arbitrary games
- Ideally we would learn the level-0 strategy from the data

Level-0 Model: Linear4

Definition (Linear4 level-0 model [Wright & Leyton-Brown 2014,2019])

$\pi_{i,0}$ is a **linear4 level-0 model** with parameters $\vec{w}$ if

$$\pi_{i,0}(a_i) \propto \sum_{f \in \mathcal{F}} w_f f(a_i),$$

where $\mathcal{F} = \{f^{\text{maxmax}}, f^{\text{maxmin}}, f^{\text{eff}}, f^{\text{fair}}, f^{\text{unif}}\}$ and

- $f^{\text{maxmax}}(a_i) = 1$ iff $a_i \in \arg\max_{a'_i \in A_i} \max_{a'_{-i} \in A_{-i}} u_i(a')$
- $f^{\text{maxmin}}(a_i) = 1$ iff $a_i \in \arg\max_{a'_i \in A_i} \min_{a'_{-i} \in A_{-i}} u_i(a')$
- $f^{\text{eff}}(a_i) = 1$ iff $a_i \in \arg\max_{a'_i \in A_i} \max_{a'_{-i} \in A_{-i}} \sum_{j \in N} u_j(a')$
- $f^{\text{fair}}(a_i) = 1$ iff $a_i \in \arg\min_{a'_i \in A_i} \min_{a'_{-i} \in A_{-i}} \max_{j,j' \in N} |u_j(a') - u_{j'}(a')|$
- $f^{\text{unif}}(a_i) = 1$ for all a_i

Summary

- Standard game theoretic solution concepts are often a poor **description** of human behavior
- Behavioral game theory attempts to induce good predictive models of human behavior in games
- These models are often parameterized
- Fitting the parameters can be treated as supervised learning exercise (next lecture)
- We considered the simplest possible case:
  - Normal form games
  - No learning/repetition
  - Simple, cognitively-inspired models