The Bargaining Game Revisited: An Interdisciplinary Framework Integrating Theory, Computation, and Experiment

Author: Runqi Li
Instructor: Professor Luyao Zhang
COMSCI/ECON 206
Duke Kunshan University

1 Introduction

Bargaining games are central to understanding how individuals and agents negotiate under limited information and strategic uncertainty. This project takes an **interdisciplinary approach**, combining **economic theory**, **computational modeling**, and **behavioral experimentation**. By comparing theoretical predictions with real-world and simulated outcomes, we explore the gap between rational choice and observed human and AI behavior.

2 Objective

- To analyze the **theoretical foundations** of bargaining and fairness.
- To construct **computational simulations** that capture strategic decision-making.
- To empirically evaluate deviations between **model predictions and behavioral results**.

3 Methodology

Computational Modeling

A simulation was developed to reproduce multi-round bargaining under variable information and strategy sets. Reinforcement learning techniques were applied to simulate adaptive agents optimizing their offers and responses. Comparative analysis between static (theoretical) and dynamic (learned) strategies illustrates how bounded rationality and adaptive feedback shape equilibrium outcomes.

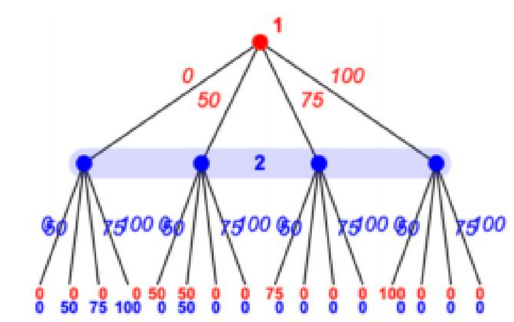
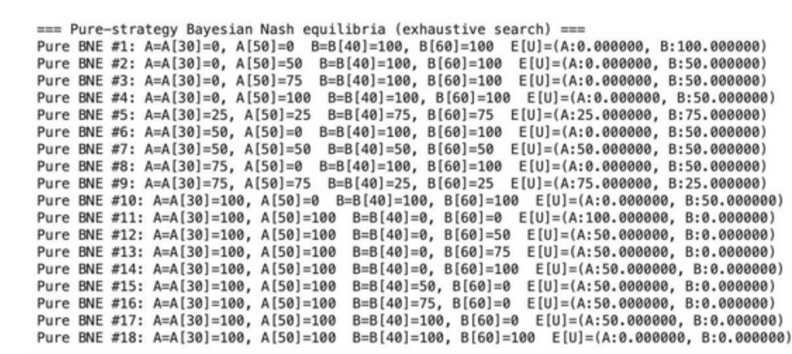


Figure 3: Extensive form tree for the sequential bargaining game



```
=== Pure-strategy Bayesian Nash equilibria (exhaustive search) ===
Pure BNE #1: A=A[30]=0, A[50]=0  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:100.000000)
Pure BNE #2: A=A[30]=0, A[50]=50  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #3: A=A[30]=0, A[50]=75  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #4: A=A[30]=0, A[50]=100  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #5: A=A[30]=25, A[50]=25  B=B[40]=75, B[60]=75  E[U]=(A:25.000000, B:75.000000)
Pure BNE #6: A=A[30]=50, A[50]=0  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #7: A=A[30]=50, A[50]=50  B=B[40]=50, B[60]=50  E[U]=(A:50.000000, B:50.000000)
Pure BNE #8: A=A[30]=75, A[50]=0  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #9: A=A[30]=75, A[50]=75  B=B[40]=25, B[60]=25  E[U]=(A:75.000000, B:25.000000)
Pure BNE #10: A=A[30]=100, A[50]=0  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:50.000000)
Pure BNE #11: A=A[30]=100, A[50]=100  B=B[40]=0, B[60]=0  E[U]=(A:100.000000, B:0.000000)
Pure BNE #12: A=A[30]=100, A[50]=100  B=B[40]=0, B[60]=50  E[U]=(A:50.000000, B:0.000000)
Pure BNE #13: A=A[30]=100, A[50]=100  B=B[40]=0, B[60]=75  E[U]=(A:50.000000, B:0.000000)
Pure BNE #14: A=A[30]=100, A[50]=100  B=B[40]=0, B[60]=100  E[U]=(A:50.000000, B:0.000000)
Pure BNE #15: A=A[30]=100, A[50]=100  B=B[40]=50, B[60]=0  E[U]=(A:50.000000, B:0.000000)
Pure BNE #16: A=A[30]=100, A[50]=100  B=B[40]=75, B[60]=0  E[U]=(A:50.000000, B:0.000000)
Pure BNE #17: A=A[30]=100, A[50]=100  B=B[40]=100, B[60]=0  E[U]=(A:50.000000, B:0.000000)
Pure BNE #18: A=A[30]=100, A[50]=100  B=B[40]=100, B[60]=100  E[U]=(A:0.000000, B:0.000000)
```

Figure 2: Solver output from NashPy (figure 2).

Behavioral Experiment

We conducted a controlled experiment involving human participants and AI counterparts in sequential bargaining tasks. Participants' behavior was recorded across rounds, measuring offer rates, acceptance thresholds, and perceived fairness. The experiment aimed to evaluate how cognitive biases and social preferences alter negotiation efficiency and stability relative to computational predictions.

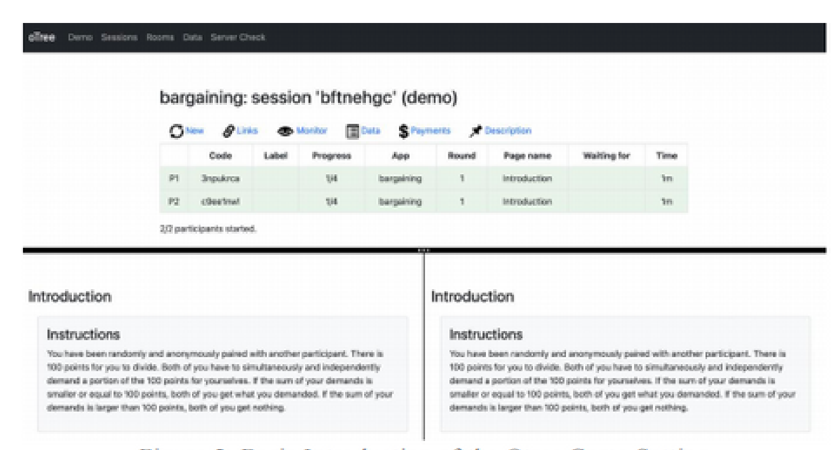
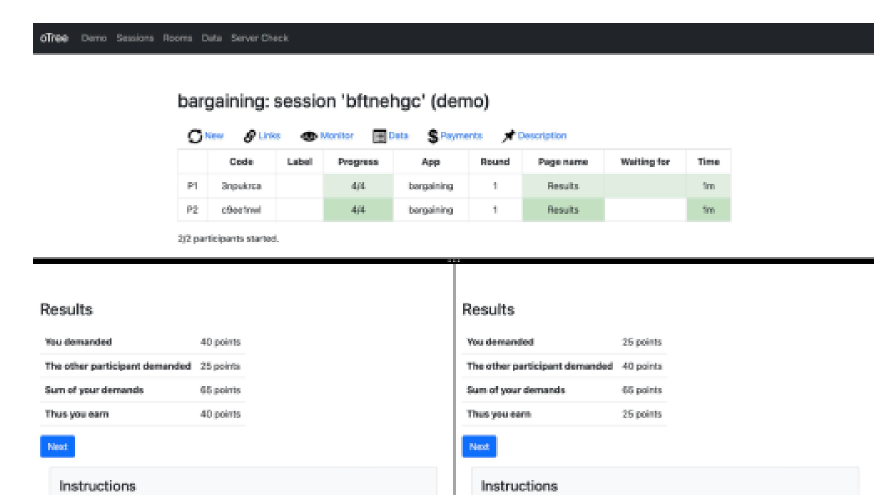
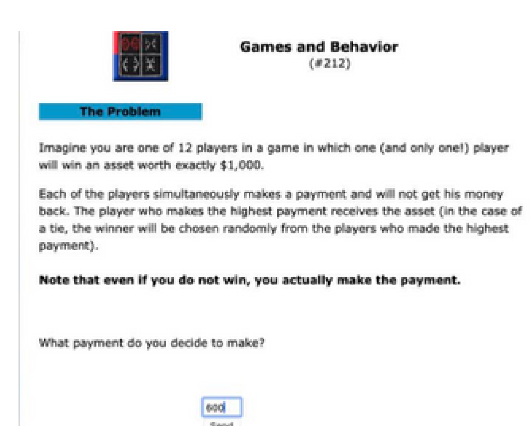


Figure 5: Basic Introduction of the Otree Game Setting



Auction Design:

An all-pay auction framework was chosen to study strategic bidding under uncertainty. Two settings were implemented: a fixed-value control ($1,000) and a stochastic-value treatment ($800–$1,200). Both human participants and AI agents (ChatGPT and Doubao) acted as bidders, always paying their bid regardless of outcome. This setup allows direct observation of the winner's curse phenomenon predicted by mechanism design theory.



AI Agent	Group	Avg Bid	Winner's Curse Rate	Avg Payoff
ChatGPT	Control	320	10%	580
ChatGPT	Treatment	360	40%	540
Doubao	Control	450	20%	520
Doubao	Treatment	480	50%	480
DeepSeek	Control	400	20%	550
DeepSeek	Treatment	430	50%	510
Qwen	Control	370	10%	570
Qwen	Treatment	410	40%	530

Table2: Experiment data for four AI agents under two conditions.

4 Results

- Google Colab simulations produced expected-value payoff matrices consistent with Bayesian structures, indicating rational responses under type uncertainty.
- Game Theory Explorer (GTE) generated subgame-perfect equilibria for sequential versions of the game, highlighting how perfect observability simplifies multi-equilibrium uncertainty.
- Human experiments (oTree) showed that participants gravitated toward fair, moderate strategies (around 50), even as resource caps scaled (10, 100, 1000).
- LLM experiments revealed framing sensitivity: fairness-oriented prompts yielded 50–50 splits, while profit-maximizing frames pushed bids toward 100.

5 Innovation

1. Game Design Innovation

The classic simultaneous demand bargaining game was extended into a Bayesian framework with **private player types representing minimum acceptable demands.** This modification transforms a **deterministic** game into a **belief-driven optimization** problem, mirroring real-world negotiations with information asymmetry. By systematically varying the resource cap (10, 100, 1000), the study evaluates how **scale** affects equilibrium efficiency and fairness.

2. Multi-Method Integration

This project uniquely **integrates analytical theory, computational modeling, behavioral experiments, and LLM simulations** within one unified bargaining framework.
Using NashPy (Colab) for BNE computation, GTE for SPNE comparison, oTree for human interaction, and LLM prompting for AI decision-making, it establishes a **multi-layered experimental ecosystem** bridging economics, computation, and cognitive science.

If AI systems can learn to act "fairly" through prompts or reward design, should we consider that genuine fairness — or merely a simulated behavioral alignment with human norms?

6 SDG Contribution

My project contributes most directly to **SDG 10: Reduced Inequalities within and among countries** by exploring how game-theoretic mechanism design, enhanced by AI simulation, can promote fairer and more inclusive resource allocation. By combining theoretical modeling with agent-based experiments, the research also aligns with **SDG 16: Peace, Justice and Strong Institutions**, as it aims to support decision systems that are not only efficient but also socially just and transparent.

References

Osborne, M. J., & Rubinstein, A. (1994). A Course in Game Theory. MIT Press.
Shoham, Y., & Leyton-Brown, K. (2008). Multiagent Systems. Cambridge University Press.
Knight, V. (2021). NashPy: A Python library for the computation of equilibria of 2-player strategic games, Version 0.0.28.
Savani, R., & von Stengel, B. (2015). Game Theory Explorer – Software for the Applied Game Theorist. Computational Management Science, 12, 5–33.
Sargent, T. J., & Stachurski, J. (2021). Quantitative Economics (Python), Version 0.5.1.