

Аннотация

Применение мультиагентного обучения с подкреплением для динамических игр двух лиц

Гимишян Ашот

Работа посвящена исследованию применения мультиагентного обучения с подкреплением в динамических играх двух лиц. В ней представлена модифицированная задача о сделке — экспериментальная игра из теории игр. Данная игра может быть рассмотрена как модель многих реальных ситуаций, где стороны ведут переговоры о разделе общего ресурса. Она иллюстрирует сложность подобных переговоров и важность учета различных факторов, включая текущую и будущую выгоду, риск, а также влияние на соперника. В ходе исследования была разработана модель искусственного интеллекта, которая путем симуляции двусторонних переговоров эффективно решает данную задачу. Проведенные эксперименты подтвердили практическую значимость модели, отраженную в монотонном возрастании средней награды и уровня согласия между участниками. Полученные результаты могут быть применены для моделирования переговоров и достижения взаимоприемлемых результатов.

Ключевые слова: Искусственный интеллект · Машинное обучение · Нейронные сети · Обучение с подкреплением · Теория игр · Динамические игры · Модифицированная задача о сделке · Переговоры

Оглавление

1	Введение	4
1.1	Описание динамических игр двух лиц и мультиагентного обучения с подкреплением	4
1.2	Обоснование выбора темы и актуальность проблемы	6
1.3	Цель работы и задачи, которые необходимо решить	7
2	Обзор литературы	9
2.1	Краткий обзор основных публикаций по теме игр двух лиц и мультиагентного обучения с подкреплением	9
2.2	Описание существующих подходов к решению задачи о сделке и их ограничения	11
2.2.1	Задача о сделке	11
2.2.2	Решение двухходовой задачи о сделке	13
2.2.3	Переговоры о цене недвижимости	13
2.2.4	Двухходовка с сокращением прибыли	14
2.2.5	Переговоры о цене недвижимости (продолжение)	15
2.2.6	Анализ бесконечно повторяющихся игр	16
3	Модифицированная задача о сделке	17
3.1	Формулировка	17
3.1.1	Общее описание	17
3.1.2	Коэффициент дисконтирования	18
3.1.3	Сила и влияние на других игроков	19
3.1.4	Динамика раундов и принятие решений	20
3.1.5	Формальное описание	20
3.1.6	Множества возможных стратегии	22
3.2	Описание предполагаемых особенностей и ограничений	22
3.2.1	Недостатки классической версии	22
3.2.2	Достоинства модифицированной версии	22

3.2.3	Основные характеристики	23
3.3	Мотивация для применения мультиагентного обучения с подкреплением в данной задаче	24
4	Методы мультиагентного обучения с подкреплением	26
4.1	Описание выбранных алгоритмов MARL	26
4.1.1	PPO	26
4.1.2	MADDPG	27
4.1.3	Q-learning	29
4.2	Обоснование выбора определенного алгоритма для данной задачи	30
4.3	Описание процесса обучения и оптимизации	31
4.3.1	Программная реализация	31
4.3.2	Математическая реализация	33
5	Результаты экспериментов	35
5.1	Экспериментальная среда	35
5.1.1	Описание	35
5.1.2	Набор данных	36
5.1.3	Оценка производительности	36
5.2	Подведение итогов экспериментов	37
5.2.1	Результаты обучения моделей	37
5.2.2	Анализ производительности	37
6	Заключение	38
6.1	Подведение итогов исследования	38
6.2	Оценка практической значимости результатов	38
6.3	Перспективы развития и дальнейшие исследования в данной области . .	39
7	Приложение	42
7.1	Реализация классов ActorNetwork и CriticNetwork на Python	42
7.2	Функция main() для инициализации параметров и запуска обучения . . .	44

Список сокращений

DCC Deep Continuous Clustering.

DDPG Deep Deterministic Policy Gradient.

DQN Deep Q-Network.

DRL Deep Reinforcement Learning.

MADDPG Multi-Agent Deep Deterministic Policy Gradient.

MARL Multi-Agent Reinforcement Learning.

PPO Proximal Policy Optimization.

RL Reinforcement Learning.

TD-learning Temporal-Difference Learning.

WoLF Win or Learn Fast.

ИИ Искусственный интеллект.

ИНС Искусственная нейронная сеть.

РН Равновесие Нэша.

Глава 1

Введение

Давайте перейдем из эпохи противостояния в эпоху переговоров...

Ричард Никсон

В современном мире существует множество задач, которые могут быть решены при помощи технологий искусственного интеллекта. Одной из таких задач является моделирование игр, в которых принимают участие несколько агентов. Это может быть полезным, например, в бизнесе, где агентами могут выступать различные компании. Для них моделирование игр может служить в качестве инструмента прогнозирования рынка, определения стратегий конкуренции и разработки маркетинговой политики. Помимо этого, моделирование игр с несколькими агентами может быть полезным во многих других областях, таких как политика и технологии. В политике моделирование игр может помочь анализировать стратегии и взаимодействия между различными участниками политических процессов. Оно может использоваться для анализа выборов, где различные партии и кандидаты могут выбирать свои стратегии в зависимости от действий своих конкурентов. В технологиях моделирование игр может быть полезно для анализа конкуренции на рынке и разработки новых продуктов и услуг. Компании могут моделировать игры для определения оптимальной стратегии ценообразования на основе поведения конкурентов.

1.1 Описание динамических игр двух лиц и мультиагентного обучения с подкреплением

Пусть у нас есть два игрока, которых мы обозначим как A и B . Каждый игрок имеет набор возможных стратегий, которые мы обозначим как S_A и S_B соответственно. Игрок A выбирает стратегию $s_A \in S_A$, а игрок B выбирает стратегию $s_B \in S_B$. Выбор стратегии каждым игроком зависит от предыдущих ходов обоих игроков. Выигрыш каждого игрока определяется функцией выигрыша, которую мы обозначим как $U_A(s_A, s_B)$ для игрока A и $U_B(s_A, s_B)$ для игрока B . Эти функции выигрыша зависят от

выбранных стратегий обоих игроков. Цель каждого игрока — выбрать такую последовательность стратегий, которая максимизирует его ожидаемый выигрыш. То есть, игрок A стремится максимизировать $U_A(s_A, s_B)$ по всем возможным s_A , а игрок B стремится максимизировать $U_B(s_A, s_B)$ по всем возможным s_B .

Что касается MARL, то это подход в машинном обучении при котором агенты обучаются путем взаимодействия друг с другом и с окружающей средой. Каждый агент получает обратную связь от среды в виде награды (англ. reward) или штрафа (англ. penalty) за каждое свое действие. Награда представляет собой меру успеха выполненного действия, а штраф — меру неудачи.

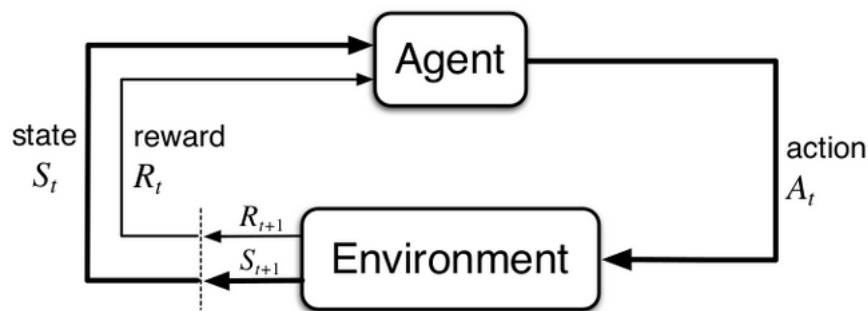


Рис. 1.1: Мультиагентное обучение с подкреплением

Для моделирования динамических игр с применением обучения с подкреплением и последующего поиска их решений необходимо создать алгоритмы, которые обеспечат взаимодействие агентов в соответствии с заданными правилами и стратегиями. Процесс моделирования можно начать с изучения конфликтной ситуации, которую планируется представить в виде игры и определения предпочтений заинтересованных сторон. В качестве следующего этапа можно рассматривать разработку стратегий и правил. Агенты будут использовать стратегий для достижения своих целей, а правила будут регулировать их поведение. Помимо этого, для эффективного моделирования игр с несколькими агентами необходимо учитывать различные факторы, такие как тип игры, количество и характеристики агентов, возможные исходы при выборе агентами той или иной стратегий. Важно также учитывать возможные изменения в игре и адаптировать стратегии и правила соответствующим образом.

1.2 Обоснование выбора темы и актуальность проблемы

Применение мультиагентного обучения с подкреплением для динамических игр двух лиц является актуальной и перспективной темой для исследования. Приведу несколько аргументов, обосновывающих выбор этой темы и актуальность проблемы:

1. Мультиагентное обучение с подкреплением является одним из передовых направлений в области искусственного интеллекта. Изучение применения MARL в динамических играх двух лиц может привести к новым открытиям в разрешении конфликтных ситуации путем переговоров.
2. В реальных сценариях взаимодействие между агентами играет ключевую роль. Исследование применения мультиагентного обучения с подкреплением в динамических играх двух лиц может помочь разработать новые методы и подходы для эффективного сотрудничества и соперничества между агентами.
3. Исследование MARL в контексте динамических игр двух лиц может привести к новым применениям в различных отраслях, таких как финансы, экономика, транспорт, робототехника и других. Это может способствовать прогрессу в этих областях и улучшению качества жизни людей.

Мультиагентное обучение с подкреплением является относительно новым направлением исследований в области искусственного интеллекта, которое показало свой потенциал в различных задачах. Мультиагентные системы становятся все более распространенными в различных областях, например, в робототехнике, автономной навигации, финансах, игровой индустрии, бизнесе, политике и других. В таких системах возникает потребность в разработке эффективных методов обучения для агентов, которые позволят им адаптироваться к изменяющимся условиям и достигать высокой производительности в условиях взаимодействия с другими агентами и окружающей средой.

1.3 Цель работы и задачи, которые необходимо решить

Кроме описания различных алгоритмов MARL для решения динамических игр двух лиц, моей целью является разработка новой игры в теории игр — модифицированной задачи о сделке. Она будет представляться впервые как расширение классической задачи о сделке. В данной игре участники должны соревноваться в некооперативном взаимодействии с целью распределения заданных общих ресурсов между собой. Также собираюсь представить новый подход к распределению ресурсов между участниками, который учитывает влияние участника на переговорный процесс и на основе этого подхода разработать модель искусственного интеллекта, способную решить модифицированную задачу о сделке. Модель будет обучена для поиска оптимальных стратегий и обеспечения эффективной коммуникации между участниками, а также для адаптации к изменяющимся условиям и поведению всех участников.

С помощью математики и искусственного интеллекта хочется создать прикладной инструмент для решения следующих проблем:

1. Формализация завершенных конфликтных ситуации, соответствующих описанию данной игры, с целью выяснения на каком этапе стороны не смогли достичь соглашения и каковы были их убытки.
2. Формализация действующих конфликтных взаимодействий на ранних этапах развития и предложение оптимальных исходов разрешающих такие ситуаций. Доказательство того, что искусственному интеллекту можно доверять разрешение конфликтов.
3. Моделирование взаимодействия, которые потенциально могут вызвать конфликты в реальной жизни. Учитывая влияние участников на обстановку, их силы, интересы и временной фактор, предложить возможные варианты решения.

Во вступительной части **первой главы** представлена ознакомительная информация. Глава включает определения MARL и динамических игр. Далее изложены аргументы мотивирующие выбор темы и её актуальность в современных условиях. В завершении главы кратко описывается проблема, которую необходимо решить, и формулируются общие цели дипломной работы.

Во **второй главе** проведен анализ основных публикаций, посвященных теме игр двух лиц и мультиагентного обучения с подкреплением. Здесь рассмотрены ключевые исследования, описывающие различные подходы и техники, которые сформировали основу современных методов в этой области. Кроме того, дано формальное описание задачи о сделке. Дополнительно сделан подробный обзор существующих подходов к решению этой задачи..

В **третьей главе** дана формулировка модифицированной задачи о сделке. Там же обоснована необходимость перехода от классической к модифицированной версии задачи о сделке. В заключительной части главы описана мотивация применения методов машинного обучения для решения данной задачи.

Четвертая глава посвящена подробному описанию и сравнению выбранных алгоритмов мультиагентного обучения с подкреплением и их применению к модифицированной задаче о сделке. Обсуждается выбор алгоритмов, процесс обучения и оптимизации, а также обосновывается их применимость для решения данной задачи.

Пятая глава представляет результаты экспериментов, проведенных для оценки эффективности разработанной модели в решении модифицированной задачи о сделке. Здесь описывается среда, в которой проводились эксперименты, и набор данных, использованный для обучения и оценки модели.

Результаты работы обсуждаются в заключительной, **шестой главе**. В ней также раскрывается прикладная значимость и ценность проведенного исследования. К тому же, данная глава содержит рекомендации по дальнейшему развитию и расширению сферы применения полученных результатов.

Глава 2

Обзор литературы

В этой главе представлю итоги анализа наиболее влиятельных и значимых результатов, связанных с играми двух лиц и мультиагентным обучением с подкреплением. Моя цель состоит в том, чтобы представить вам ключевые исследования, которые сформировали основные направления и подходы в данной области. Я уделю внимание на разнообразие методов, применяемых учеными для решения задач и преодоления сложностей, связанных с мультиагентными системами. Этот обзор позволит лучше понять существующие подходы и их взаимосвязь, а также определить перспективы развития данной области науки. Кроме того, я дам формальное описание классической задачи о сделке. Далее в работе будет представлено формальное описание модифицированной задачи о сделке, поэтому важно уже сейчас ознакомиться с классической версией этой экспериментальной игры, чтобы понять различия.

2.1 Краткий обзор основных публикаций по теме игр двух лиц и мультиагентного обучения с подкреплением

В своей статье [1] Джон Нэш предложил понятие РН, которое стало фундаментальным в теории игр. РН представляет собой набор стратегий в игре для двух и более игроков, в котором ни один участник не может увеличить выигрыш, изменив свою стратегию, если другие участники своих стратегий не меняют. Нэш доказал существование такого равновесия в смешанных стратегиях в любой конечной игре.

Здесь [2] Майкл Литтман предложил марковские игры (стохастические игры) в качестве основы для мультиагентного обучения с подкреплением. Марковские игры являются расширением марковских процессов принятия решений, учитывающих действия нескольких агентов. Автор применил этот подход к кооперативным и конкурентным задачам, демонстрируя его применимость в различных сценариях.

Джералд Тесауро представил TD-Gammon в [3]. Это программа для игры в нар-

ды, использующую **TD-обучение** (обучение с подкреплением на основе временной разницы). TD-Gammon стал одним из первых успешных примеров применения обучения с подкреплением в играх и продемонстрировал возможность разработки высокопроизводительных агентов без явных знаний. Модель была обучена без знания правил и стратегий игры, а только на основе информации об успехе или неудаче своих ходов. Такой подход показал возможность создания высокопроизводительных агентов без предварительного знания о специфике проблемы, что стало важным шагом в развитии обучения с подкреплением и искусственного интеллекта в целом.

В своей работе [4] Боулинг и Велозо представили алгоритм для мультиагентного обучения с использованием переменной скорости обучения (**WoLF**). Алгоритм меняет скорость обучения в зависимости от оценки эффективности текущей стратегии агента. Этот подход позволяет агентам лучше адаптироваться к изменяющимся ситуациям в мультиагентных средах.

Обзор [5] представляет собой широкий анализ методов мультиагентного обучения с подкреплением, включая их теоретические основы, алгоритмы и применение в различных областях. Авторы обсуждают как кооперативные, так и конкурентные сценарии, а также коммуникацию между агентами. Они также рассматривают вопросы сходимости, обучения и адаптации в мультиагентных системах.

В статье [6] авторы представили метод глубокого мультиагентного обучения с подкреплением, который позволяет агентам обучаться совместной кооперативной коммуникации с использованием **DCC**. Это подход к обучению, в котором агенты обмениваются информацией и настраивают свои стратегии на основе обратной связи других агентов, что позволяет им успешно решать сложные задачи, требующие кооперации.

В работе [7] авторы представили алгоритм **MADDPG**, который является расширением алгоритма DDPG для мультиагентных сред. MADDPG основан на актор-критике, где каждый агент имеет собственные актор и критик, обучаемые независимо. Он позволяет агентам совместно обучаться в смешанных средах, где они могут взаимодействовать как кооперативно, так и конкурентно. Этот алгоритм продемонстрировал успех в решении сложных мультиагентных задач, таких как управление движением и командные игры.

Статья [8] исследует применение мультиагентного обучения с подкреплением, используя **Q-обучение** в ультиматумной игре. Авторы применяют Q-обучение для обучения агентов и используют ϵ -жадную стратегию, позволяющую агентам исследовать и эксплуатировать среду. Цель исследования состоит в том, чтобы анализировать и

сравнивать различные стратегии агентов, чтобы определить наиболее эффективные подходы в ультиматумной игре.

В статье [9] автор применяет глубокое обучение с подкреплением (**DRL**) для обучения агентов в мультиагентных переговорах. Это исследование направлено на поиск оптимальных стратегий и результатов в сложных переговорах, где участники обсуждают несколько вопросов. Агенты, представленные глубокими нейронными сетями, учатся оптимальной стратегии, основываясь на взаимодействии с окружающей средой и получении подкрепления за свои действия.

[10] — статья, в которой представлен метод **PPO**. Данный метод призван решить некоторые проблемы с обучением с подкреплением, связанные с нестабильностью и большим временем обучения. PPO является методом оптимизации политики, который использует технику, известную как обрезка (англ. clipping), чтобы ограничить обновления политики в каждой итерации оптимизации.

Работа [11] является одной из первых, в которой обсуждается сочетание глубокого обучения и обучения с подкреплением. Авторы представили алгоритм, который они назвали **DQN**. Он сочетает Q-обучение, метод обучения с подкреплением, с глубокими нейронными сетями.

В 4-ой главе подробно описаны некоторые из упомянутых алгоритмов. Там же выбран алгоритм для решения модифицированной задачи о сделке. Подбор подходящего алгоритма сопровождается подробными комментариями.

2.2 Описание существующих подходов к решению задачи о сделке и их ограничения

2.2.1 Задача о сделке

Задача о сделке — экспериментальная экономическая игра двух лиц в теории игр, которая моделирует процесс двусторонних переговоров. В этой игре участники должны договориться о распределении конечного объема общих ресурсов. На практике такое взаимодействие возникает в случае, когда торговля может принести прибыль, например, одна из сторон ценит ресурсы меньше. Более того, для моделирования переговоров и анализа такой проблемы необходимо отсутствие установленных ценовых норм на рынке.

Итак, на столе переговоров денежная сумма M — прибыль от торговли. Игроки ходят поочередно. Суть игры заключается в том, что первый игрок предлагает распределить M в пропорциях $(x, M - x)$. Такой дележ реализуется при наличии согласия второго игрока. В результате его реализации первый получает x , второй — $M - x$. Если игроки не приходят к соглашению, то каждый получает свою часть заранее фиксированного исхода (a, b) . Первый получает a , второй — b . Следует отметить, что в подавляющем большинстве реальных случаев $a = b = 0$. Предполагается, что $a + b < M$.

Формально задачу о сделке можно представить в виде кортежа (N, X, u) , где

- N — множество игроков. В этой игре $|N| = 2$, то есть участвуют два игрока, которые обычно называются Игрок 1 и Игрок 2.
- X — пространство исходов. $X = \{(x, y) \in \mathbb{R}_{\geq 0}^2 : x + y = M\}$. Оно является непрерывным и линейным. В этой работе мы не будем нормировать его, то есть не будем рассматривать $\frac{x}{M} + \frac{y}{M} = 1$ вместо $x + y = M$.
- $u = (u_1, u_2)$ — пара функций полезности, которая представляет предпочтения каждого игрока относительно результатов в X . Функция полезности $u_i : X \rightarrow \mathbb{R}_{\geq 0}$ представляет предпочтения Игрока i .

Цель игры — добиться соглашения между участниками о том, как распределить ресурсы. Из-за разногласий в предпочтениях по результатам в X можно наблюдать конфликт интересов. Задача о сделке моделирует этот конфликт как игру стратегического взаимодействия, где каждый игрок пытается договориться о сделке, которая максимизирует его собственную функцию полезности.

Данная задача является важной для понимания того, как агенты принимают решения в условиях конкуренции и ограниченности ресурсов. Ее формализация помогает лучше понять проблемы ведения переговоров в различных экономических ситуациях. Были разработаны различные методы решения этой игры. Одним из распространенных способов решения является решение Нэша. Оно представляет собой набор результатов, которые оба игрока предпочитают любому альтернативному результату. Решение Нэша основано на концепции справедливости, которая предполагает, что оба игрока имеют равную силу в переговорах.

2.2.2 Решение двухходовой задачи о сделке

Игра «Ультиматум» — двухходовая задача о сделке, описанная выше, является простейшим вариантом задачи о сделке, которая имеет следующее решение: второй игрок примет предложение $(x, M - x)$, если $M - x \geq b$. Предположим, $x = M - b$. Если первый игрок изменит свою стратегию с целью максимизации собственного выигрыша, то второму будет выгодно отказать. Кроме того, заметим, что при получении предложения $(M - b, b)$ любая стратегия (принять или отклонить) второго игрока приносит прибыль не больше b . Следовательно, исход $(M - b, b)$ является РН. Здесь очевидным недостатком является наличие равновесия $(M, 0)$ при $b = 0$. Таким образом, в реальной ситуации первый агент может оставить себе всю денежную сумму M . Формально это будет считаться равновесием.

2.2.3 Переговоры о цене недвижимости

Пример из реальной жизни:

- Продавец не готов продать свою недвижимость меньше 4,500,000 рублей.
- Покупатель готов платить не больше 5,000,000 рублей.
- $M = 5,000,000 - 4,500,000 = 500,000, a = b = 0$.
- Предположим продавец ходит первым и обладает совершенной информацией. Он знает, что покупатель отклонит любое предложение $> 5,000,000$ и примет любое $\leq 5,000,000$.
- Продавец максимизирует свою прибыль предлагая цену равную 5,000,000 или $x = 500,000$. Покупатель принимает, ведь $M - x \geq b$.
- Продавец получает 500,000, то есть всю денежную сумму M , покупатель — 0.

Вывод: покупатель не должен заранее объявить максимальную сумму, которую готов платить за недвижимость. Эту информацию продавец может использовать против него. В том числе из этих соображений я буду рассматривать игру с неполной информацией.



Рис. 2.1: Переговоры о стоимости недвижимости

2.2.4 Двухходовка с сокращением прибыли

Предположим, что $M = 12$, $\delta = \frac{1}{3}$ — коэффициент дисконтирования. Обратите внимание, что во втором раунде $M = 4$. Игрок 2 знает, что может получить 3.99 в раунде 2, поскольку Игрок 1 предпочтет 0.01 ничему и примет дележ (0.01, 3.99). Это следует из предположения что игроки ориентированы на будущее, рациональны и максимизируют свой выигрыш. Учитывая это, Игрок 1 должен предложить Игроку 2 $4.00 > 3.99$ в первом раунде, оставив себе $12 - 4 = 8$. Поскольку Игрок 2 понимает, что ни в каком исходе не может получить больше 4.00, он примет дележ (8.00, 4.00). Предложение Игрока 1 отдать 4.00 Игроку 2 в первом раунде равно сумме ставки в начале финального раунда, умноженной на коэффициент дисконтирования, $\delta \cdot M$. Такое предложение приведет к равновесию. Это обобщается на любую задачу о сделке с n раундами, где $2 < n < \infty$.

Этот пример показывает, что в игровой ситуации, где участники направлены на долгосрочные результаты, действуют рационально и стараются максимизировать свои выигрыши, возможно достижение РН, которое обеспечивает оптимальные исходы. В данном случае, для достижения РН Игрок 1 в первом раунде предлагает Игроку 2 условия, которые выгодны обоим, чтобы Игрок 2 принял предложение и не стремился получить больше в следующем раунде.

Вывод: в задачах о сделке можно достичь РН, которое приводит к эффективному результату, если участники сделки обладают достаточной информацией о возможных исходах, являются рациональными и ориентированы на будущее. Однако, если эти условия не выполняются, могут возникнуть проблемы координации и нежелательные исходы.

2.2.5 Переговоры о цене недвижимости (продолжение)

- Минимальная стоимость, по которой продавец продаст свою недвижимость составляет 15,000,000 рублей, а максимальная цена, которую покупатель готов заплатить — 16,000,000 рублей. Следовательно, $M = 1,000,000$ рублей.
- Оба игрока имеют одинаковый коэффициент дисконтирования $\delta = 0.8$.
- Процесс переговоров ограничен двумя раундами. Это обусловлено тем, что продавец обязан продать недвижимость до определенного срока (возможно, для покупки другой недвижимости), а покупателю, в свою очередь, нужно приобрести эту недвижимость до определенной даты.
- В первом раунде переговоров предложение выдвигает покупатель, а во втором раунде — продавец.
- Следует применить метод обратной индукции, начав анализ с последнего, второго раунда переговоров и двигаться в обратном направлении для поиска оптимальной последовательности действий.
- С точки зрения сегодняшнего дня потенциальный суммарный выигрыш от сделки во втором раунде составляет $\delta \cdot M$. На этом этапе на продавце лежит обязанность предложить контрпредложение.
- В заключительном раунде продавец предложит оставить себе $\delta \cdot M$, предоставляя покупателю возможность принять или отклонить это предложение. На этом этапе покупателю нет разницы между принятием или отклонением предложения. В обоих случаях его выигрыш равен 0.
- Осознавая это, покупатель в первом раунде должен предложить продавцу $\delta \cdot M$. Продавец, будучи равнодушным между ожиданием и немедленным принятием, принимает данное предложение.
- В нашем примере, где $\delta = 0.8$ и $M = 1,000,000$, покупатель предлагает продавцу $0.8 \cdot M$ или 800,000, оставляя себе 200,000. Таким образом, стоимость недвижимости составляет $15,000,000 + 800,000 = 15,800,000$.

2.2.6 Анализ бесконечно повторяющихся игр

- Теперь предположим, что количество переговорных раундов не ограничено. Переговоры могут продолжаться бесконечно.
- Если покупатель делает предложение первым, то сумма $x(1) \cdot M$, которую он намерен оставить себе в первом раунде, должна гарантировать продавцу приемлемую выгоду. Эта выгода должна быть не меньше той, которую продавец может получить в следующем раунде, если отклонит текущее предложение и предложит забрать $y(2) \cdot M$. Иными словами, в первом раунде продавец должен получить сумму, эквивалентную $\delta \cdot y(2) \cdot M$.
- Покупатель предлагает продавцу $(1 - x)M = \delta yM$, и таким образом, $x = 1 - \delta y$.
- Аналогичным образом, продавец должен предложить покупателю $(1 - y)M = \delta xM$. Следовательно, $y = 1 - \delta x$.
- Получаем систему уравнений:

$$\begin{cases} x &= 1 - \delta(1 - \delta x), \\ y &= 1 - \delta(1 - \delta y). \end{cases} \quad (2.1)$$

- $x = y = \frac{1-\delta}{1-\delta^2}$. Обратите внимание, что $x + y > 1$.
- x и y обозначают суммы, которые получают покупатель и продавец соответственно, если они делают первое предложение в первом раунде.

Практические выводы

1. В реальной жизни стороны переговоров не знают коэффициенты дисконтирования друг друга или их относительные уровни терпения, но могут пытаться угадать эти значения.
2. Нужно подать сигнал о том, что вы терпеливы, даже если на самом деле нет. Например, не отвечать контрпредложениями сразу же.
3. Более терпеливый игрок получает большую часть суммы M , которая находится на столе переговоров.