



ФЭН НИУ ВШЭ

31 мая 2024 г.

Моделирование конкурентных стратегий гетерогенных фирм с помощью обучения с подкреплением

Выполнил: Перепелкин Владимир Андреевич БЭК205

Научный руководитель: Пильник Николай Петрович

Агентное моделирование: позволяет из описания взаимодействий на микроуровне понять, как будет развиваться вся система.

Обучение с подкреплением: позволяет агентам исследовать модель (среду) и выработать такую стратегию (политику) π , которая помогает решить задачу максимизации дисконтированной награды.

Проблемы агентного моделирования в экономике:

- Недостаток микроэкономических оснований.
- Большое количество гиперпараметров.
- Задание правил поведения агентов (стратегий, политик) извне. Они могут быть неадекватными их целям в конкретной среде. Это решается эндогенизацией поведения агентов по отношению к среде с помощью обучения с подкреплением.
- Производственный сектор экономики, как правило, довольно сильно упрощается.

Цель работы: разработка агентной модели многоотраслевой экономики несовершенной конкуренции с фирмами, политика которых вырабатывается с помощью обучения с подкреплением, а также последующий анализ динамики развития агентной модели и поведения фирм.

Рынок — то, посредством чего фирмы взаимодействуют друг с другом.

Фирма — единственный вид агента в модели.

Константа (Гиперпараметр)
Переменная состояния
Обучаемая функция

Характеристики рынка	Обозначение	Пояснение
Количество фирм	n	
Количество отраслей	k	
Матрица цен	$P \in \mathbb{R}^{n \times k}$	$P_{i,j}$ — цена на j у фирмы i
Матрица объёмов	$V \in \mathbb{R}^{n \times k}$	$V_{i,j}$ — объёмы j у фирмы i

Характеристика фирмы	Обозначение
Производственная функция	$f_{prod}(\vec{x}_{in}, \mathcal{K}) = \min(f_{prod}^*(\vec{x}_{in}), \mathcal{K}) \rightarrow \vec{x}_{out}$
Инвестиционная функция	$f_{invest}(\vec{x}_{in}) \rightarrow \Delta \mathcal{K}$
Срок жизни основного капитала	d
Деньги	m
Запасы (резервы)	$r\vec{es}$
Основной капитал (лимит)	\mathcal{K}
Политика	$\pi_{\theta}(\text{observation}) \rightarrow \text{actions}$

где $\vec{x}_{in} \in \mathbb{R}^k$ — вектор входа. $\vec{x}_{out} \in \mathbb{R}^k$ — вектор выхода.






Ход одной фирмы:

Получает наблюдение. Фирма η получает наблюдение $o_\eta = (P, V, \vec{r}es_\eta, m_\eta, K_\eta)$. Информации о $m, \vec{r}es, K$ других фирм у неё нет.

Получает параметры действий. Фирма применяет свою текущую политику к наблюдению и семплирует параметры действий:

$$\pi_\eta(o_\eta) \rightarrow (A^{buy}, \bar{a}^{prod}, \bar{a}^{invest}, \bar{a}^{sale}, \bar{a}^{prices})$$

Действует:

-  **Покупает товары на рынке,** пополняя запасы. Доля денег, которая тратится на товар j фирмы i : $A_{i,j}^{buy} \in [0, 1]$.
 -  **Инвестирует.** Доля запаса j , идущая на основной капитал: $a_j^{invest} \in [0, 1]$.
 -  **Производит.** Доля запаса j , идущая на производство: $a_j^{prod} \in [0, 1]$
 -  **Выставляет товары на рынок.** Доля запаса j , идущая на продажу: $a_j^{sale} \in [0, 1]$
 -  **Назначает цены на свои товары.** Цена на j : $a_j \cdot (p_{\max} - p_{\min}) + p_{\min} \in [p_{\min}, p_{\max}]$
- Амортизация.** Капитал, приобретенный d периодов назад, пропадает.

Алгоритм: Multi-Agent Proximal Policy Optimzation



Используется для обучения функций π_1, \dots, π_n , задающих политику фирм.



Децентрализованные акторы $\pi_1(o_1; \theta_1), \dots, \pi_n(o_n; \theta_n)$. Задача каждого актора i :

$$\pi_i \approx \arg \max_{\pi_i} \mathbb{E}_{\tau \sim \pi_1 \dots \pi_i \dots \pi_n} \left[\sum_{t=0}^{\infty} \gamma^t r_{i,t} \right]$$

где

$\tau \sim \pi_1, \dots, \pi_n$ — траектория системы (s_0, s_1, \dots) при стратегиях π_1, \dots, π_n

$r_t = \mathbb{E}[R_t(s_t, a_t \sim \pi(s_t))]$ — ожидание награды на момент t , если действия фирмы a_t берутся из её текущей стратегии π_i

s_t — состояние системы на момент t



Централизованный критик $V(o_1, \dots, o_n; \phi)$. Предоставляет акторам оценку ожидаемой дисконтированной награды, получаемой при текущих стратегиях:

$$\hat{V}_i^{\pi_1, \dots, \pi_n}(s) \approx \mathbb{E}_{\tau \sim \pi_1 \dots \pi_i \dots \pi_n} \left[\sum_{t=0}^{\infty} \gamma^t r_{i,t} \mid s_0 = s \right]$$

Функция награды (полезности, выигрыша)

- Критик учится по состоянию, в котором находится каждая фирма, оценивать её ожидаемую дисконтированную награду $V_i^{\pi_1, \dots, \pi_n}(s)$
- Каждая фирма i вырабатывает стратегию π_i , которая максимизирует оценку ожидаемой дисконтированной награды (полезности):

$$\pi_i \approx \arg \max_{\pi_i} \frac{1}{|\mathcal{B}| \cdot |\tau|} \sum_{j=0}^{|\mathcal{B}|} \sum_{t=0}^{|\tau_j|} \hat{V}_i^{\pi_1, \dots, \pi_n}(s_{t, \tau_j})$$

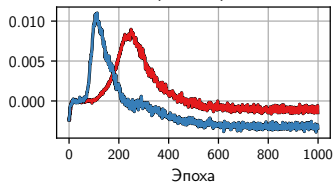
где s_{t, τ_j} — состояние системы на траектории τ_j в момент t

- Проводятся эксперименты с разными функциями наград

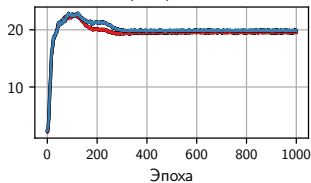
Тип награды	Функция награды
Производственная	$R_t^{prod} = \log \left(\text{const} + \sum_{j=1}^k f^{prod}(\vec{x}_{in})_j \right)$
Финансовая	$R_t^{fin} = \text{revenue}_{t+1} - \text{costs}_t$
Смешанная	$R_t^{mix} = R_t^{fin} + \nu \cdot R_t^{prod}$

- **Награда — производственная.** Задача фирм — максимизация производства.
- Производственные функции фирм — симметричные. Первая фирма производит 2 товара x_2 из одного товара x_1 .
- Инвестиционные функции фирм — одинаковые. x_1 и x_2 создают 2 единицы \mathcal{K} .

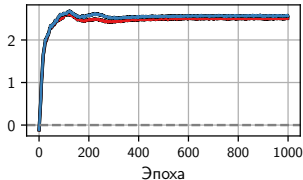
Потери актора \mathcal{L}^A



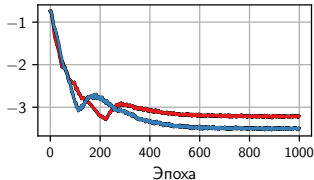
Потери критика \mathcal{L}^C



Средняя награда



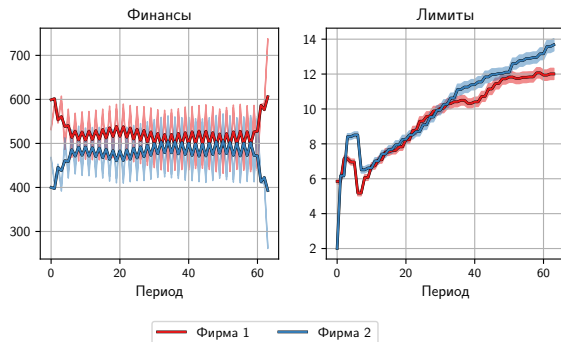
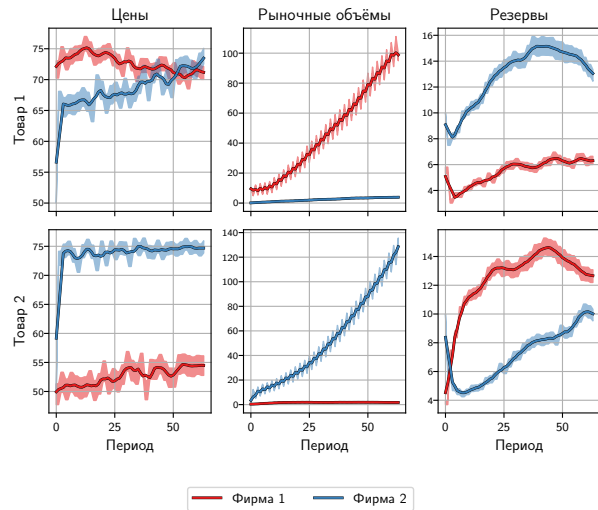
Энтропийный штраф



— Фирма 1 — Фирма 2

Фирма	f^{prod}	f^{invest}
1	$(2x_2, 0)$	$2 \min(x_1, x_2)$
2	$(0, 2x_1)$	$2 \min(x_1, x_2)$

Параметр	Значение
d (срок жизни \mathcal{K})	2
Начальные лимиты	2
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{prod}
Тип чисел	Целые



- Производство растёт, лимиты тоже растут
- Фирмы торгуют на большие суммы
- Фирмы держат довольно высокие цены на свою продукцию

Рис.: Параметры действий фирмы 1

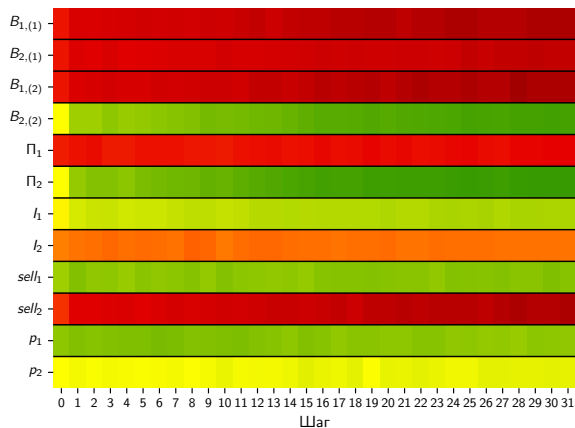
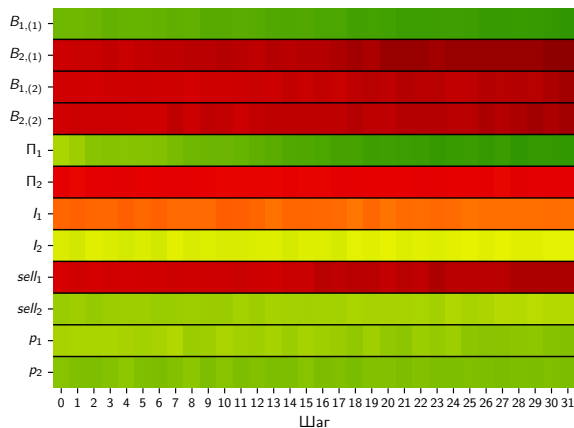
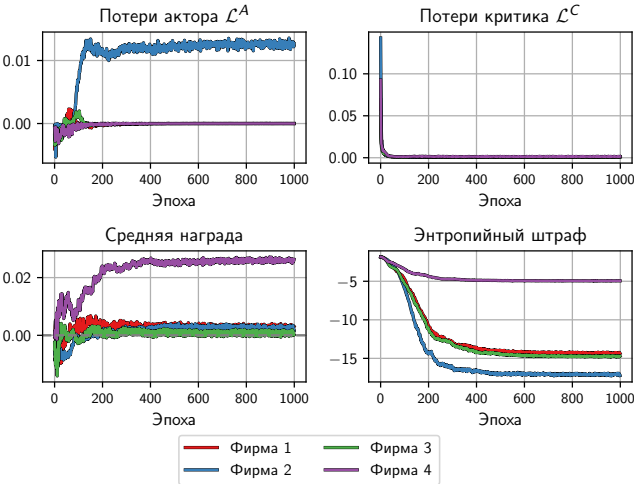


Рис.: Параметры действий фирмы 2



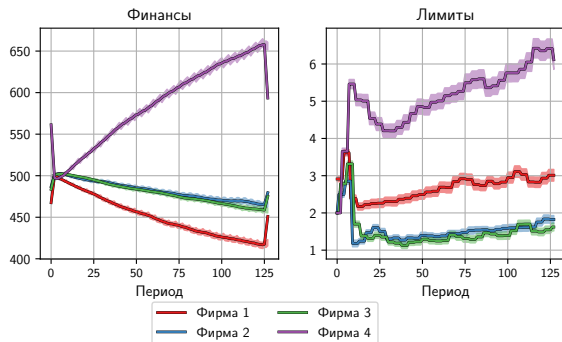
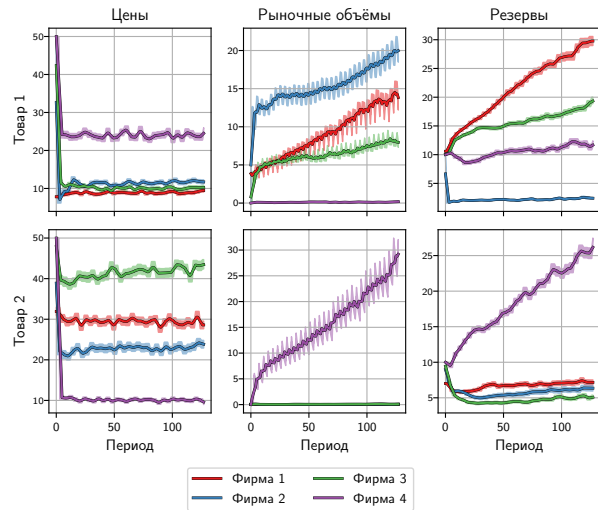
$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

- **Награда — смешанная.** Фирмам важны и прибыль, и производство.
- Производственные функции фирм — симметричные. Первая фирма производит 2 товара x_2 из одного товара x_1 . Инвестиционные функции фирм — одинаковые.
- **Фирма 4 — монополия.**



Фирма	f^{prod}	f^{invest}
1	$(2x_2, 0)$	$2 \min(x_1, x_2)$
2	$(2x_2, 0)$	$2 \min(x_1, x_2)$
3	$(2x_2, 0)$	$2 \min(x_1, x_2)$
4	$(0, 2x_1)$	$2 \min(x_1, x_2)$

Параметр	Значение
d (срок жизни \mathcal{K})	2
Начальные лимиты	2
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{mix}
Тип чисел	Целые



- Финансы перетекают из конкурентной отрасли к монополии
- Фирмы назначают почти одинаковые цены на *производимые ими* товары

Рис.: Фирма 1

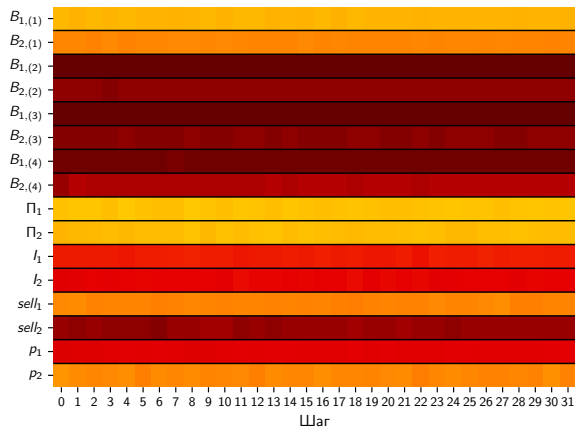
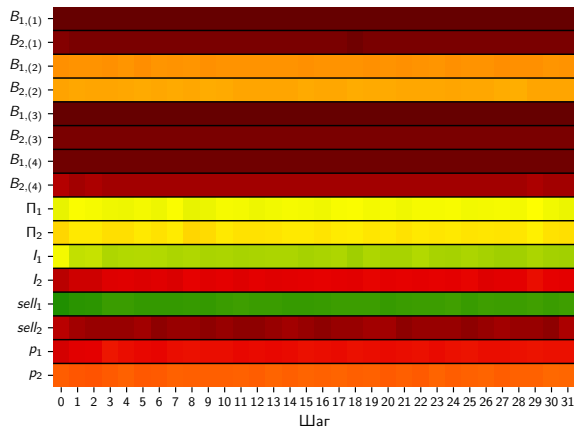


Рис.: Фирма 2



$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

Рис.: Фирма 3

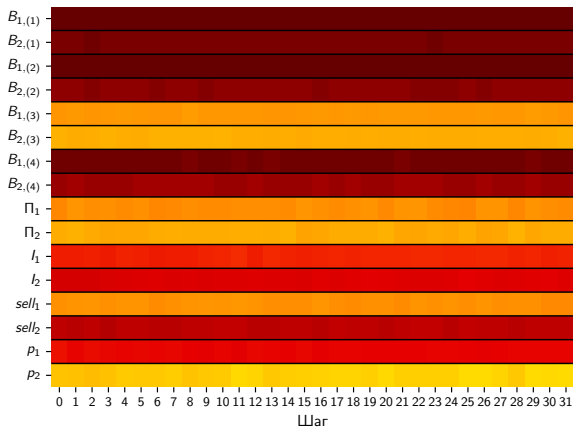
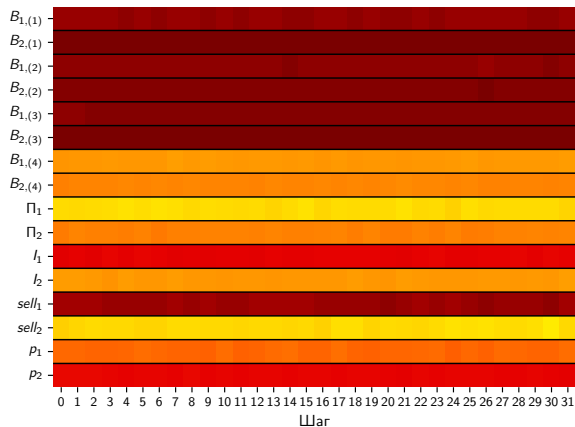


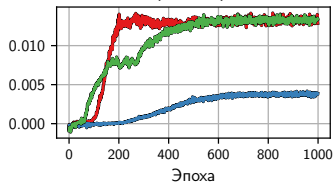
Рис.: Фирма 4 (монополия)



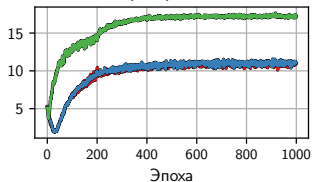
$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

- **Награда — производственная.** Фирмам важно производство.
- Производственные функции фирм — симметричные. Фирма 1 производит 3 товара x_1 из x_2 и x_3 .
- **Основной капитал создаётся только из ресурсов фирмы 3.**

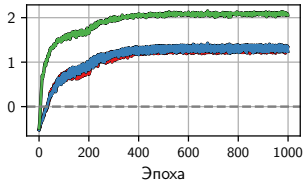
Потери актора \mathcal{L}^A



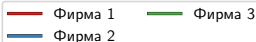
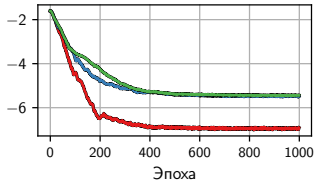
Потери критика \mathcal{L}^C



Средняя награда

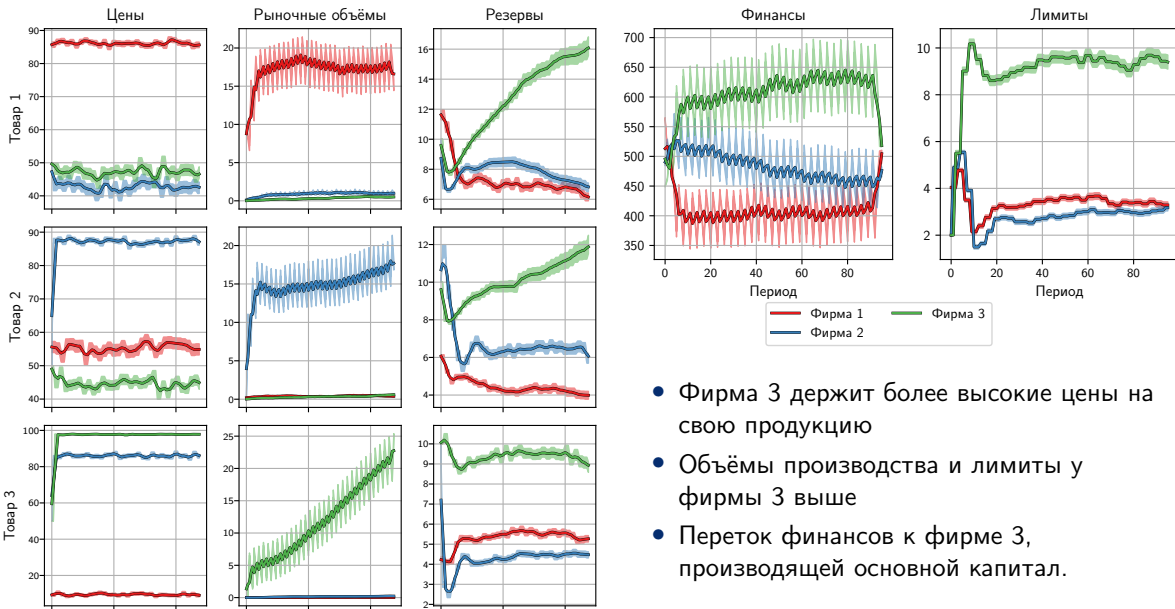


Энтропийный штраф



Фирма	f^{prod}	f^{invest}
1	$(3 \min(x_2, x_3), 0, 0)$	x_3
2	$(0, 3 \min(x_1, x_3), 0)$	x_3
3	$(0, 0, 3 \min(x_1, x_2))$	x_3

Параметр	Значение
d (срок жизни \mathcal{K})	2
Начальные лимиты	2
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{prod}
Тип чисел	Целые



- Фирма 3 держит более высокие цены на свою продукцию
- Объёмы производства и лимиты у фирмы 3 выше
- Переток финансов к фирме 3, производящей основной капитал.

Рис.: Фирма 1

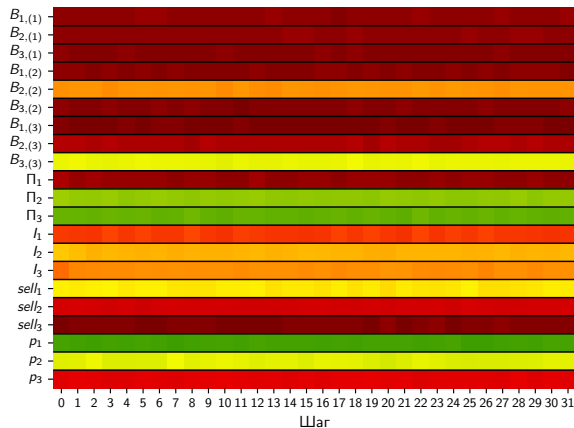
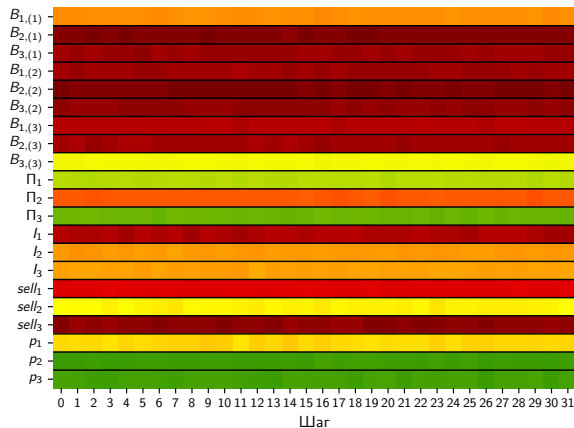
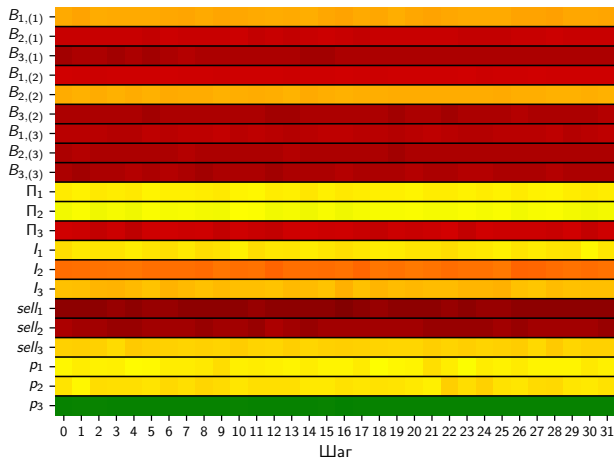


Рис.: Фирма 2



$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

Рис.: Фирма 3 (производит основной капитал)



$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

Синергетические эффекты:

- Цены отражают дефицитность согласно производственным технологиям
- Ускоренный рост отрасли, производящей ресурсы для основного капитала
- Внутриотраслевая конкуренция \Rightarrow фирмы устанавливают приблизительно одинаковые цены
- Переток денежных ресурсов от конкурентной отрасли к монополии

Стратегии фирм:

- При усреднении, как правило, меняются несильно от эпохи к эпохе \Rightarrow агрегированные параметры действий можно описать константой или простым уравнением динамики
- Без усреднения могут сильно меняться от эпохи к эпохе, поскольку политика стохастична.
- При производственной награде иногда стратегии фирм отклоняются от гипотетической идеальной стратегии, лучшей для обоих агентов. Например, повышение цен.

Будущие исследования:

- **Изменение архитектуры рынка.** Текущая архитектура рынка не подходит для большого числа агентов.
- **Добавление других видов агентов.** Государство, банки, потребители.
- **Адаптация модели для регулирования:** данный подход позволяет понять, какие политики могут возникнуть у фирм в той или иной среде \Rightarrow можно использовать для выработки оптимальных механизмов регулирования хозяйственной деятельности.



ФЭН НИУ ВШЭ

31 мая 2024 г.

Моделирование конкурентных стратегий гетерогенных фирм с помощью обучения с подкреплением

Выполнил: Перепелкин Владимир Андреевич БЭК205

Научный руководитель: Пильник Николай Петрович

-  Brusatin, S. и др. (2024). *Simulating the economic impact of rationality through reinforcement learning and agent-based modelling*. [arXiv: 2405.02161 \[cs.LG\]](#).
-  Curry, M. и др. (2022). *Analyzing Micro-Founded General Equilibrium Models with Many Agents using Deep Reinforcement Learning*. [arXiv: 2201.01163 \[cs.GT\]](#).
-  Dwarakanath, K. и др. (2024). *ABIDES-Economist: Agent-Based Simulation of Economic Systems with Learning Agents*. [arXiv: 2402.09563 \[cs.MA\]](#).
-  Leontief, W. (1986). *Input-output economics*. Oxford: Oxford University Press.
-  Lowe, R. и др. (2020). *Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments*. [arXiv: 1706.02275 \[cs.LG\]](#).
-  Schulman, J. и др. (2017). *Proximal Policy Optimization Algorithms*. [arXiv: 1707.06347 \[cs.LG\]](#).
-  Yu, C. и др. (2022). *The Surprising Effectiveness of PPO in Cooperative, Multi-Agent Games*. [arXiv: 2103.01955 \[cs.LG\]](#).
-  Леонидов, А. В. и Е. Е. Серебрянникова (2017). «Динамическая модель несовершенной конкуренции в многосекторной экономике». В: *Пробл. управл.* 4, с. 8—16. URL: <http://mi.mathnet.ru/pu1035>.

Архитектуры

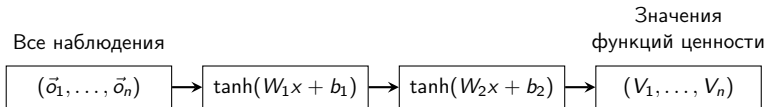


Рис.: Архитектура централизованного критика. $\phi = (W_1, b_1, W_2, b_2)$

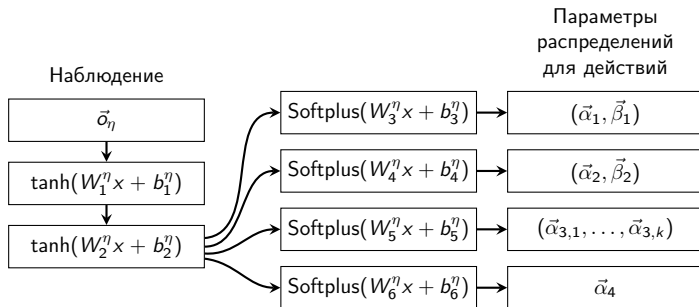


Рис.: Архитектура актора для фирмы η . $\theta_\eta = (W_1^\eta, \dots, W_6^\eta, b_1^\eta, \dots, b_6^\eta)$

Алгоритм. Получение параметров действий.

- 1: $a_j^{sale} \sim \text{Beta}(\alpha_{1,j}, \vec{\beta}_{1,j})$ ▷ Доли резервов на продажу
- 2: $a_j^{prices} \sim \text{Beta}(\alpha_{2,j}, \vec{\beta}_{2,j})$ ▷ Параметры цен
- 3: $(a_j^{prod}, a_j^{invest}, a_j^{save}) \sim \text{Dir}(\alpha_{3,j})$ ▷ Доли резервов на инвестиции и производство
- 4: $a^{buy} \sim \text{Dir}(\alpha_4)$ ▷ Доли затрат на покупки
- 5: $A_{i,j}^{buy} \leftarrow a_{i+(j-1) \cdot k}^{buy}$ ▷ Получение подходящей размерности

Алгоритм. Ход фирмы ч.1

Покупка товаров на рынке

- 1: $V_{i,j}^{bought} \leftarrow \min\{m_\eta \cdot \frac{A_{i,j}^{buy}}{P_{i,j}}, V_{i,j}\}$ ▷ Определение объёмов покупок
- 2: $m_\eta \leftarrow m_\eta - \text{tr}(P^T V^{bought})$ ▷ Уменьшение финансов фирмы
- 3: $m_i \leftarrow m_i + \langle V_i^{bought}, P_i \rangle$ ▷ Получение выручки фирмами
- 4: $\vec{res}_j^\eta \leftarrow \vec{res}_j^\eta + \sum_i^k V_{i,j}^{bought}$ ▷ Обновление резервов

Алгоритм. Ход фирмы ч.2

Инвестиции

$$5: \vec{x}_{invest} \leftarrow \vec{res} \odot a^{invest}$$

$$6: \mathcal{K}_t^\eta \leftarrow f_{invest}^\eta(\vec{x}_{invest})$$

$$7: \mathcal{K}^\eta \leftarrow \mathcal{K}^\eta + \mathcal{K}_t^\eta$$

$$8: \vec{res}^\eta \leftarrow \vec{res}^\eta - \vec{x}_{invest}$$

Производство товаров

$$9: \vec{x}_{in} \leftarrow \vec{res} \odot a_{prod}^{use}$$

$$10: \vec{x}_{out} \leftarrow f_{prod}^\eta(\vec{x}_{in})$$

$$11: \vec{res}^\eta \leftarrow \vec{res}^\eta + \vec{x}_{out} - \vec{x}_{in}$$

Выставление товаров на рынок

$$12: \vec{x}_{sale} \leftarrow \vec{res} \odot a^{sale}$$

$$13: V_\eta \leftarrow V_\eta + \vec{x}_{sale}$$

$$14: \vec{res}^\eta \leftarrow \vec{res}^\eta - \vec{x}_{sale}$$

Назначение цен и амортизация

$$15: P_\eta \leftarrow a^{prices} \cdot (p_{max} - p_{min}) + p_{min}$$

$$16: \mathcal{K}^\eta \leftarrow \sum_{\tau=t-d}^t \mathcal{K}_\tau^\eta$$

▷ Определение резервов для инвестиций

▷ Производство основного капитала

▷ Обновление лимитов

▷ Обновление резервов

▷ Определение резервов для производства

▷ Производство ресурсов

▷ Обновление резервов

▷ Определение резервов, идущих на рынок

▷ Приращение на рынке

▷ Обновление резервов

▷ Изменение цен на рынке

▷ изнашивание основного капитала

Алгоритм. Сбор траекторий

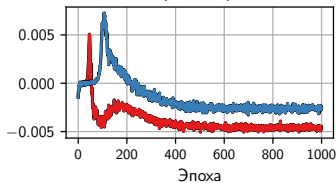
- 1: Инициализация модели
- 2: $R_{it} \leftarrow 0 \quad \forall i \in \{1, \dots, n\}, t \in \{1, \dots, T\}$
- 3: **for** $t = 1$ to T **do** ▷ Запуск модели на T периодов
- 4: **for** agent $i = 1$ to n **do** ▷ Ход каждого агента
- 5: Получение наблюдения $o_i(V, P, \vec{res}_i, m_i, \mathcal{K}_i, \mathcal{K}_{i,t-d})$
- 6: Семплирование действий $a_i \sim \pi(o_i; \theta_i)$
- 7: Получение вероятностей действий $p(a_i|o_i; \theta_i)$
- 8: Выполнение a_i , обновление среды
- 9: **if** финансовая награда **then**
- 10: $R_{t,i} \leftarrow R_{t,i} - \text{costs}_{t,i}$
- 11: $R_{t-1,i} \leftarrow R_{t-1,i} + \text{revenue}_{t,i}$
- 12: **if** производственная награда **then**
- 13: $R_{t,i} \leftarrow R_{t,i} + \nu \cdot \log \left(\text{const} + \sum_{j=1}^k f^{prod}(\vec{x}_{in})_j \right)$
- 14: $p_{t,i}, s_{t,i}, a_{t,i} \leftarrow p(a_i|o_i; \theta_i), o_i, a_i$
- 15: $s'_{t,i}, v_{t,i} \leftarrow s_{t-1,i}, V_i(s_t; \phi)$
- 16: Подсчёт $\hat{A}^{GAE}(v, R)$
- 17: $\hat{V}_{t,i} \leftarrow v_{t,i} + \hat{A}_{t,i}^{GAE}$
- 18: $\tau_{t,i} \leftarrow \{\hat{V}_{t,i}, \hat{A}_{t,i}^{GAE}, s_{t,i}, s'_{t,i}, p_{t,i}, a_{t,i}, v_{t,i}\}$ ▷ Сохранение траекторий

Алгоритм. Multi-Agent Proximal Policy Optimization

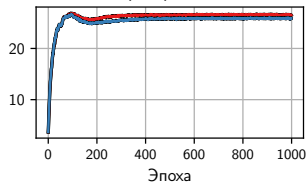
- 1: Собрать набор из K траекторий $\mathcal{D} \leftarrow (\mathcal{D}_1, \dots, \mathcal{D}_K)$
- 2: **for** случайный мини-батч b из D **do**
- 3: $s, s', a, p, v, \hat{V}, \hat{A}^{GAE} \leftarrow b$
- 4: **for** agent $i = 1$ to n **do**
- 5: $V_{t,i}^{old} \leftarrow v_{t,i}$
- 6: $a_{t,i}, p_{t,i} \leftarrow a_{t,i}, p_{t,i}$
- 7: $V_{t,i}^{clip} \leftarrow \text{clip}(V_{t,i}(s_t; \phi), V_{t,i}^{old} - 0.2, V_{t,i}^{old} + 0.2)$
- 8: $\mathcal{L}_i^C(\phi) \leftarrow \frac{1}{T} \sum_{t=1}^T \max\{L_1(V_{t,i}(s_t; \phi), \hat{V}_{t,i}), L_1(V_{t,i}^{clip}, \hat{V}_{t,i})\}$
- 9: Нормализация \hat{A}^{GAE}
- 10: Получение вероятностей действий $p(a_i|s_i; \theta_i)$ для актуальных параметров θ_i
- 11: $\text{ratios}_{t,i} \leftarrow \exp(\log p(a_i|s_i) - \log p_{old}(a_i|s_i))$
- 12: $\text{ratios}_{t,i}^{clip} \leftarrow \text{clip}(\text{ratios}_{t,i}, 0.8, 1.2)$
- 13: $\mathcal{L}_i^A(\theta_i) \leftarrow - \sum_{t=1}^T \min\{\hat{A}_{t,i}^{GAE} \cdot \text{ratios}_{t,i}, \hat{A}_{t,i}^{GAE} \cdot \text{ratios}_{t,i}^{clip}\}$
- 14: Обновление параметров актора i : $\theta_i \leftarrow \arg \min_{\theta_i} \mathcal{L}_i^A(\theta_i)$
- 15: Обновление параметров централизованного критика: $\phi \leftarrow \arg \min_{\phi} \mathcal{L}_i^C(\phi)$

- **Награда — производственная.** Фирмам важна только прибыль.
- Производственные функции фирм — симметричные. Первая фирма производит 2 товара x_2 из одного товара x_1 .
- Основной капитал бесконечный, фирмы не инвестируют.

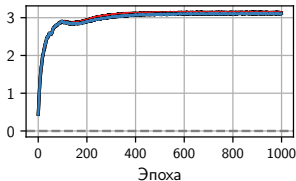
Потери актора \mathcal{L}^A



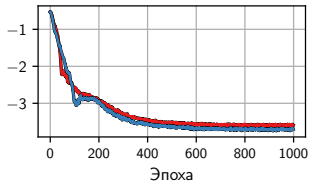
Потери критика \mathcal{L}^C



Средняя награда



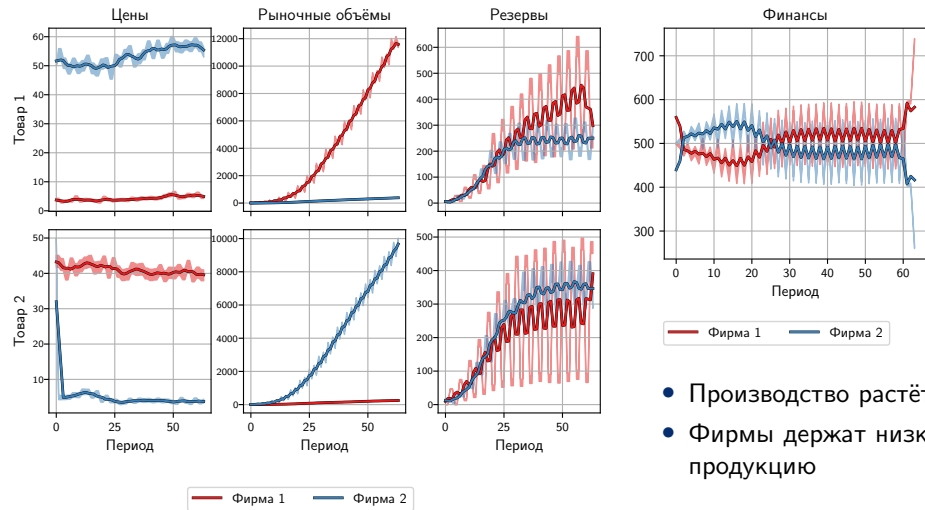
Энтропийный штраф



— Фирма 1 — Фирма 2

Фирма	f^{prod}	f^{invest}
1	$(2x_2, 0)$	$2 \min(x_1, x_2)$
2	$(0, 2x_1)$	$2 \min(x_1, x_2)$

Параметр	Значение
d (срок жизни \mathcal{K})	∞
Начальные лимиты	∞
Начальные резервы	$(10, 10)$
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{prod}
Тип чисел	Целые



- Производство растёт очень интенсивно
- Фирмы держат низкие цены на свою продукцию

Рис.: Фирма 1

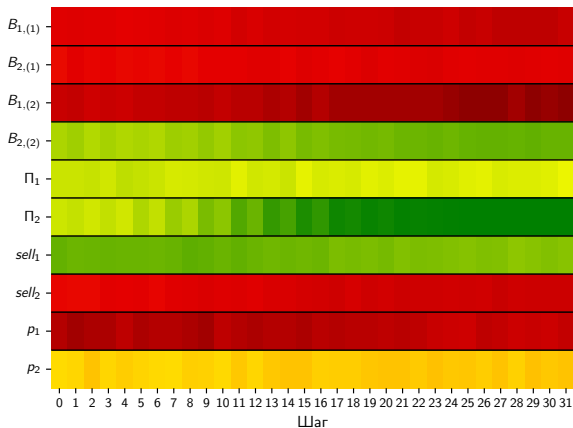
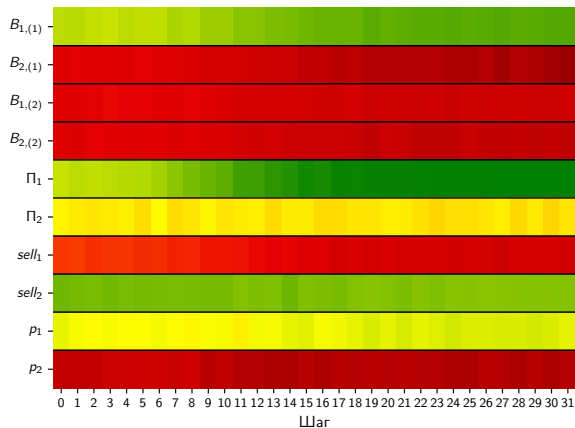
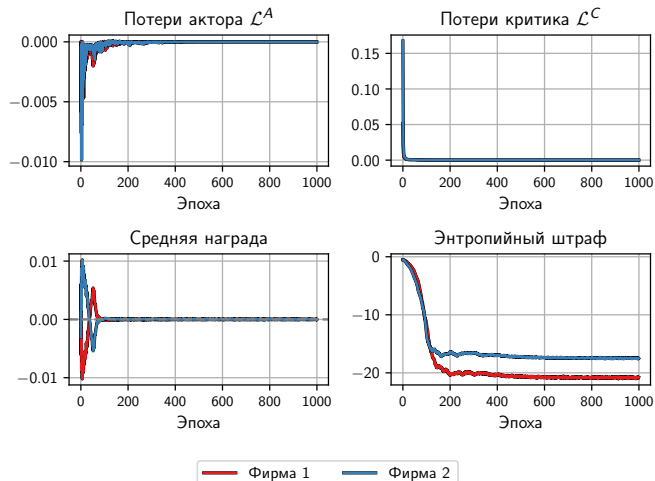


Рис.: Фирма 2



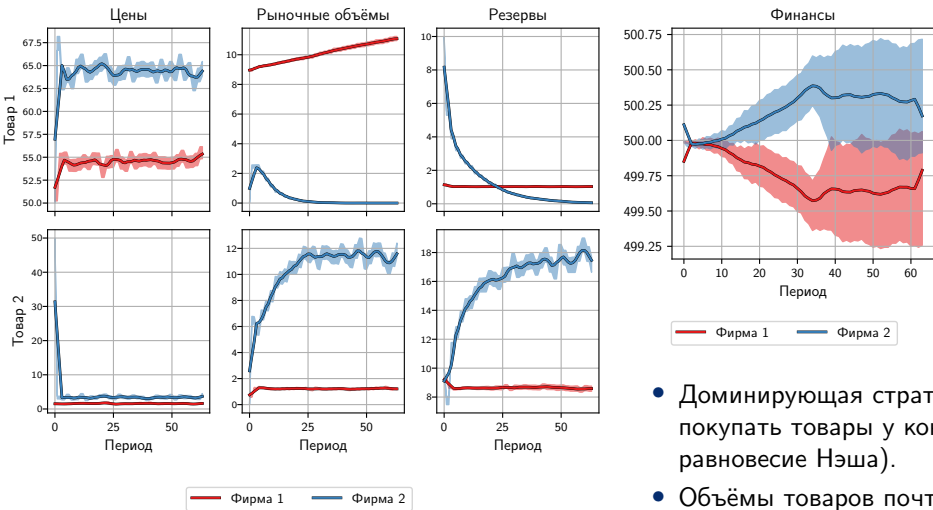
$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

- **Награда — финансовая.** Фирмам важна только прибыль.
- Производственные функции фирм — симметричные. Первая фирма производит 2 товара x_2 из одного товара x_1 .
- Основной капитал бесконечный, фирмы не инвестируют.



Фирма	f^{prod}	f^{invest}
1	$(2x_2, 0)$	$2 \min(x_1, x_2)$
2	$(0, 2x_1)$	$2 \min(x_1, x_2)$

Параметр	Значение
d (срок жизни \mathcal{K})	∞
Начальные лимиты	∞
Начальные резервы	$(10, 10)$
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{fin}
Тип чисел	Целые



- Доминирующая стратегия фирм — не покупать товары у контрагента (это равновесие Нэша).
- Объёмы товаров почти не растут, несмотря на отсутствие лимитов.

Рис.: Фирма 1

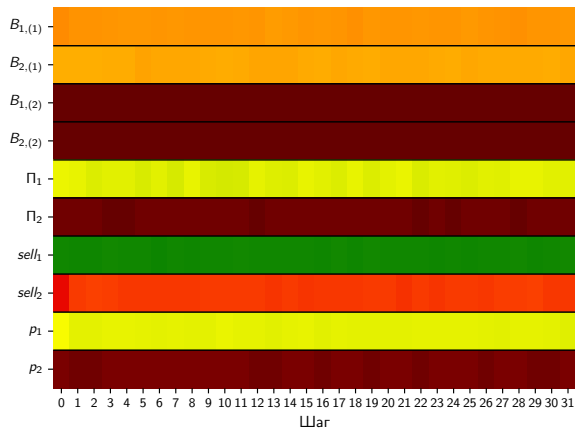
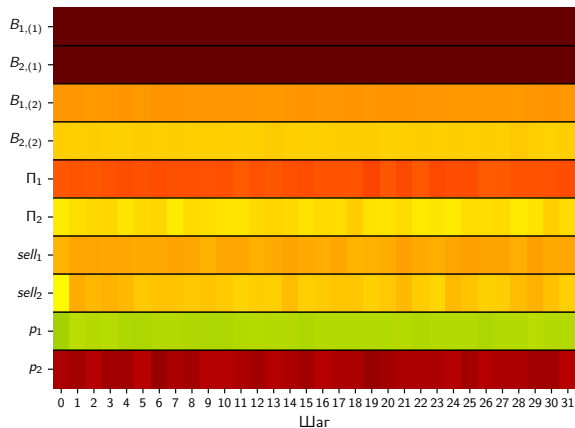


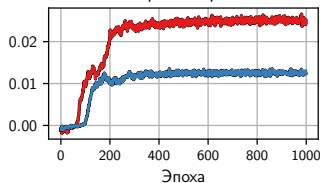
Рис.: Фирма 2



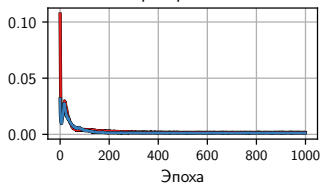
$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

- **Награда — смешанная.** Фирмам важны и прибыль, и производство.
- Производственные функции фирм — симметричные. Первая фирма производит 2 товара x_2 из одного товара x_1 .
- Инвестиционные функции фирм — одинаковые. x_1 и x_2 создают 2 единицы \mathcal{K} .

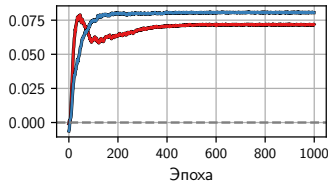
Потери актора \mathcal{L}^A



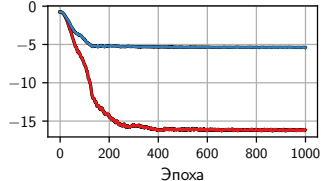
Потери критика \mathcal{L}^C



Средняя награда



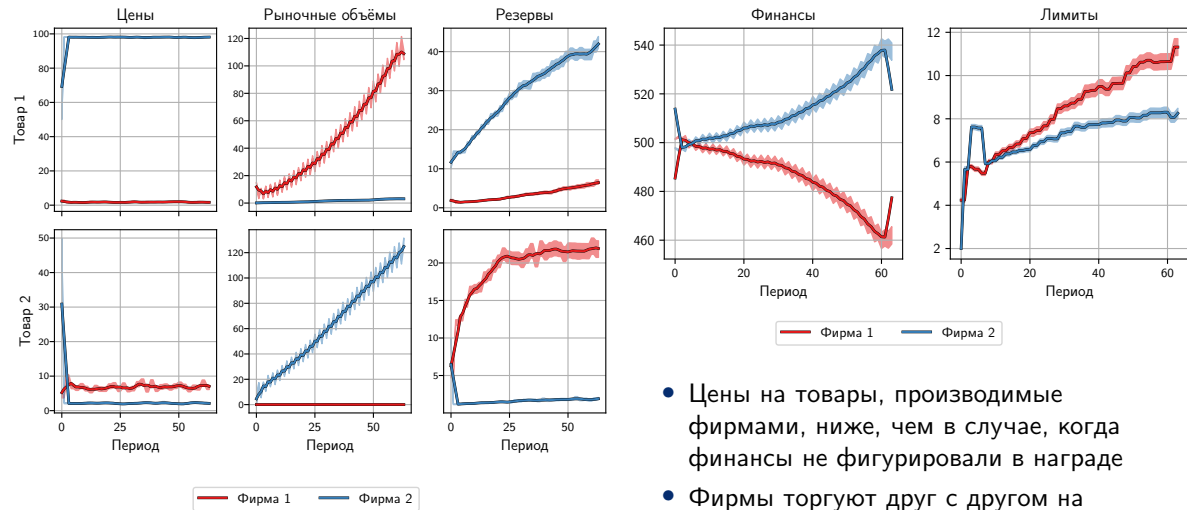
Энтропийный штраф



— Фирма 1 — Фирма 2

Фирма	f^{prod}	f^{invest}
1	$(0, 2x_1)$	$2 \min(x_1, x_2)$
2	$(2x_2, 0)$	$2 \min(x_1, x_2)$

Параметр	Значение
d (срок жизни \mathcal{K})	2
Начальные лимиты	2
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{mix}
Тип чисел	Целые



- Цены на товары, производимые фирмами, ниже, чем в случае, когда финансы не фигурировали в награде
- Фирмы торгуют друг с другом на меньшие суммы

Рис.: Фирма 1

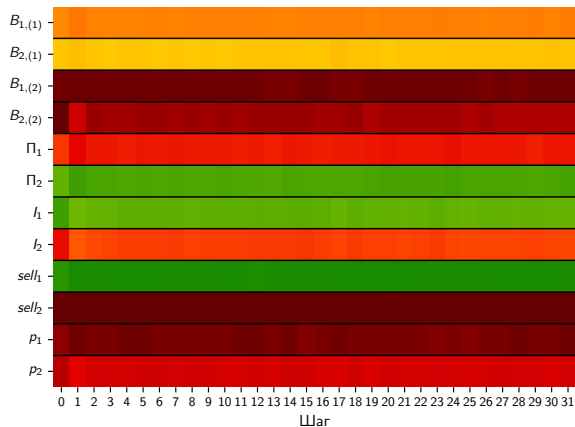
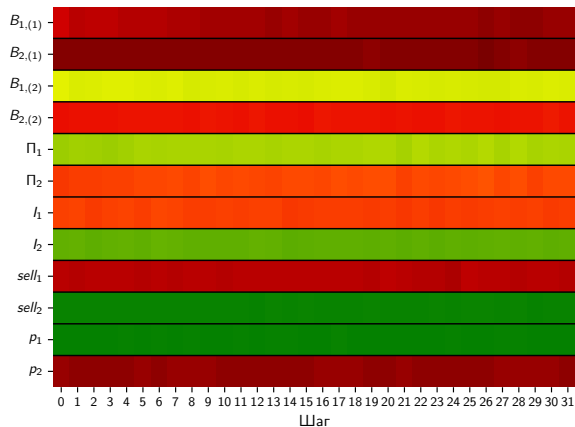


Рис.: Фирма 2



$B_{j,(i)}$	% Потрачено на j у фирмы i
Π_j	% Использовано j на производство
I_j	% Использовано j на инвестиции
$sell_j$	% Продано j

Рис.: Фирма 1

Рис.: Фирма 2

