



ФЭН НИУ ВШЭ

20 мая 2024 г.

Моделирование конкурентных стратегий гетерогенных фирм с помощью обучения с подкреплением

Выполнил: Перепелкин Владимир Андреевич БЭК205

Научный руководитель: Пильник Николай Петрович



Содержание

1. Введение
2. Обзор литературы
3. Агентная модель
4. Обучение модели
5. Эксперименты
6. Итоги
7. Вопросы и ответы



Проблемы агентного моделирования в экономике:

- Недостаток микроэкономических оснований.
- Большое количество гиперпараметров.
- Задание правил поведения агентов 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.



- **Micro-Founded General Equilibrium Models with Many Agents using Deep Reinforcement Learning** (Curry и др., 2022).

Переменные управления фирмы — зарплаты w , цены p , Δk

Производственные функции: $Y_{i,t} = A_i k_{i,t}^{1-\alpha} L_{i,t}^\alpha$

Формула для изменения капитала: $k_{i,t+1} = k_{i,t} + \Delta k_{i,t}$

- **ABIDES-Economist** (2024, февраль). (Dwarakanath и др., 2024)

Объёмы производства зависят от случайного экзогенного фактора $\varepsilon_{t,j}$

$$\epsilon_{t,j} = (\epsilon_{t-1,j})^{\rho_j} \exp(\varepsilon_{t,j})$$

$$y_{t,j} = \epsilon_{t,j} (L_{t,j})^{\alpha_j}$$

- **Rational macro ABM** (2024, май). (Brusatin и др., 2024)

Два типа агентов: рациональные и следующие тренду.

Две отрасли: средства производства и предметы потребления.

Переменные управления фирм — $Y_{i,t}^*$, $P_{i,t}$.

Производственные функции: $Y_{i,t} = \min(\alpha_N N_{i,t}, \alpha_K K_{i,t})$

Предметы потребления не хранятся на рынке больше одного периода.



Агентная модель

Рынок и фирмы

Характеристики рынка	Обозначение
Количество фирм	n
Количество отраслей	k
Матрица цен	$P \in \mathbb{R}^{n \times k}$
Матрица объёмов товаров	$V \in \mathbb{R}^{n \times k}$

Характеристика фирмы	Обозначение
Денежные ресурсы	m
Срок жизни основного капитала	d
Основной капитал (лимит)	$\mathcal{K} = \sum_{\tau=t-d}^t \mathcal{K}_{\tau}$
Производственная функция	$f_{prod}(\vec{x}_{in}, \mathcal{K}) =$ $= \min(f_{prod}(\vec{x}_{in}), \mathcal{K}) \rightarrow \vec{x}_{out}$
Инвестиционная функция	$f_{invest}(\vec{x}_{in}) \rightarrow \mathcal{K}_{new}$
Политика	$\pi(\text{observation}) \rightarrow \text{actions}$



1. **Получение наблюдения.** Фирма η получает наблюдение $o_\eta = (P, V, \vec{res}_\eta, m_\eta, \mathcal{K}_\eta)$, соответствующее текущему состоянию среды.
2. **Получение параметров действий.** Фирма применяет свою текущую политику к наблюдению и семплирует параметры действий:

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

3. **Действия.**

Алгоритм. Ход фирмы ч.1

Покупка товаров на рынке

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

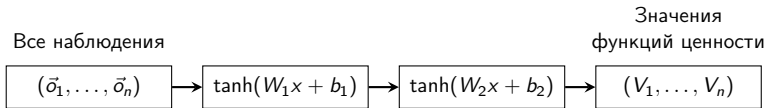


Рис.: Архитектура централизованного критика. $\phi = (W_1, b_1, W_2, b_2)$

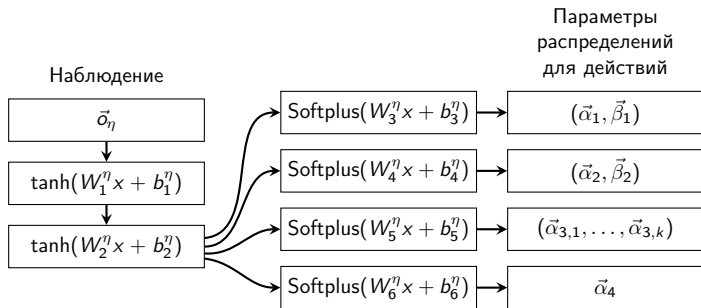


Рис.: Архитектура актора для фирмы η . $\theta_\eta = (W_1^\eta, \dots, W_6^\eta, b_1^\eta, \dots, b_6^\eta)$



Алгоритм — Multi-Agent Proximal Policy Optimization (Schulman и др., 2017, Yu и др., 2022, Lowe и др., 2020).

- Централизованный критик $V(o_1, \dots, o_n; \phi)$.
- Децентрализованные акторы $\pi_1(o_1; \theta_1), \dots, \pi_n(o_n; \theta_n)$.

Функция награды. Задача агента — максимизация $\mathbb{E}(\sum_{t=1}^{\infty} r_t \gamma^{t-1})$.

- Производственная

$$R_t^{prod} = \log \left(\epsilon + \sum_{j=1}^k f^{prod}(\vec{x}_{in})_j \right)$$

- Финансовая

$$R_t^{fin} = \text{revenue}_{t+1} - \text{costs}_t$$

- Смешанная

$$R_t^{mix} = R_t^{fin} + \nu \cdot R_t^{prod}$$



Эксперимент 1.

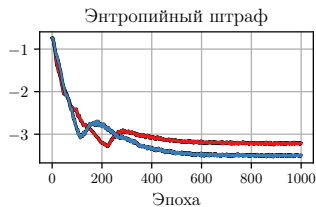
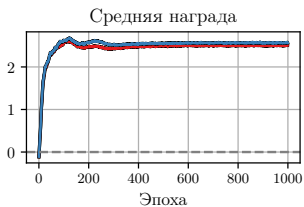
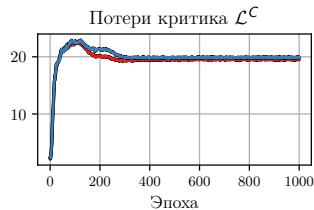
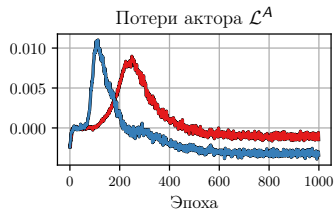
Обучение модели. «Производство, 2 симметричные отрасли»

Таблица: Функции фирм в эксперименте 1

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

Таблица: Параметры эксперимента 1

Параметр	Значение
d (срок жизни 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 данных	int64



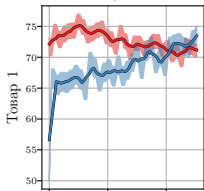
— Фирма 1 — Фирма 2



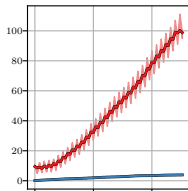
Эксперимент 1.

Динамика среды. «Производство, 2 симметричные отрасли»

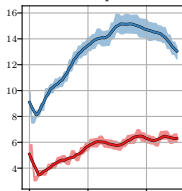
Цены



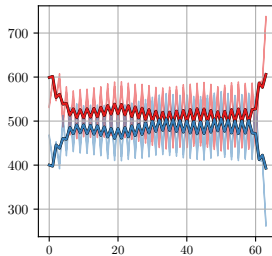
Рыночные объёмы



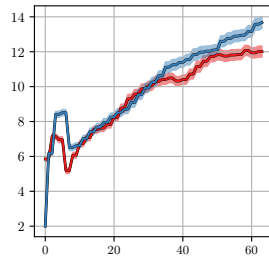
Резервы



Финансы



Лимиты



— Фирма 1 — Фирма 2

Период

Период

Период

— Фирма 1 — Фирма 2



Эксперимент 1.

Политика фирм. «Производство, 2 симметричные отрасли»

Рис.: Параметры действий фирмы 1

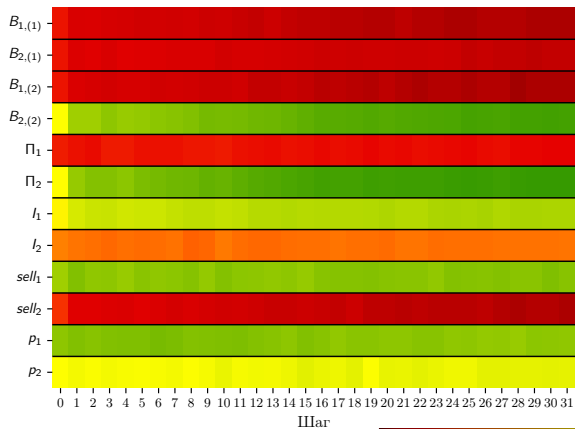
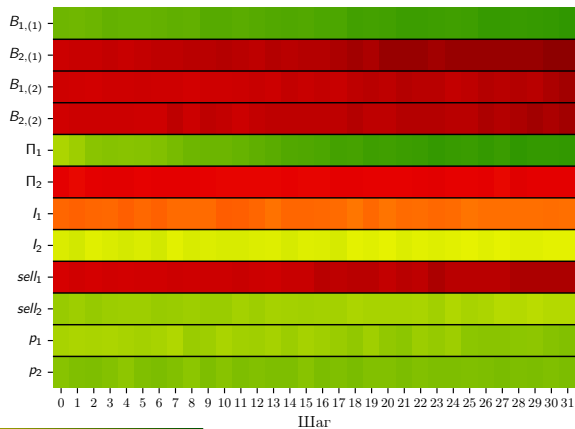


Рис.: Параметры действий фирмы 2





Эксперимент 2.

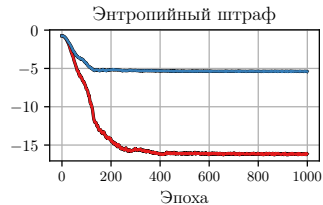
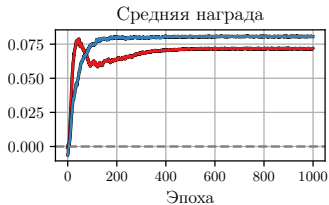
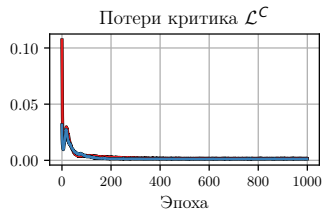
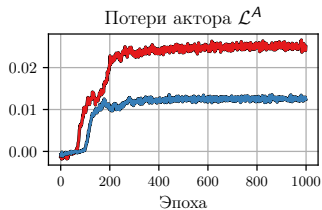
Обучение модели. «Производство и финансы, 2 симметричные отрасли»

Таблица: Функции фирм в эксперименте 2

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

Таблица: Параметры эксперимента 2

Параметр	Значение
d (срок жизни 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 данных	int64

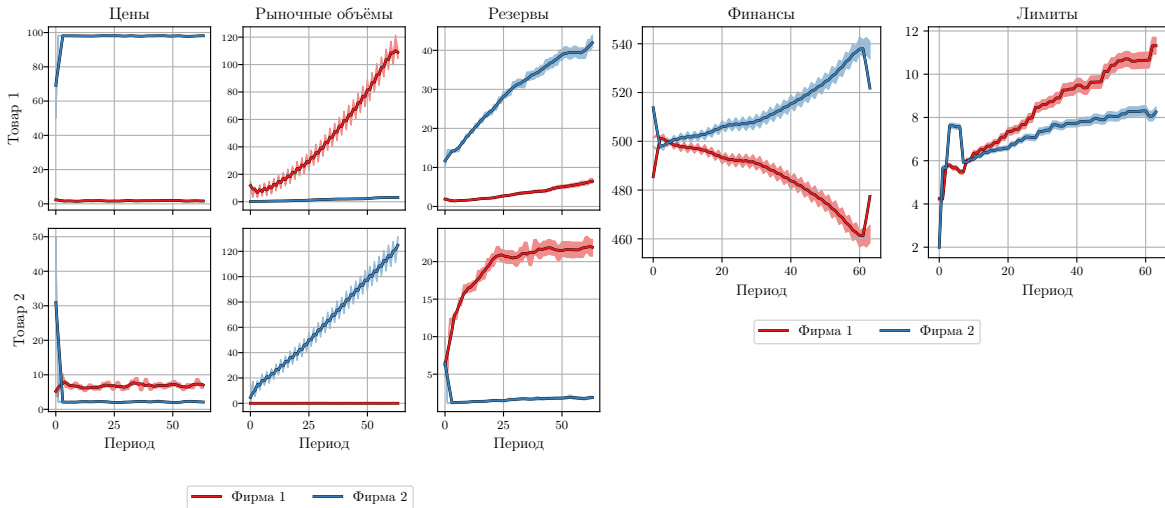


— Firma 1 — Firma 2



Эксперимент 2.

Динамика среды. «Производство и финансы, 2 симметричные отрасли»





Эксперимент 2.

Политика фирм. «Производство и финансы, 2 симметричные отрасли»

Рис.: Параметры действий фирмы 1

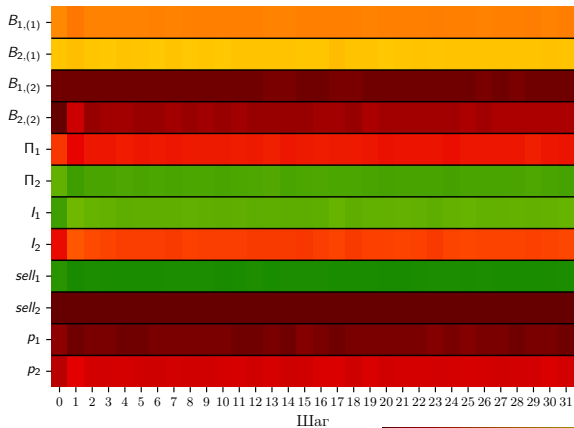
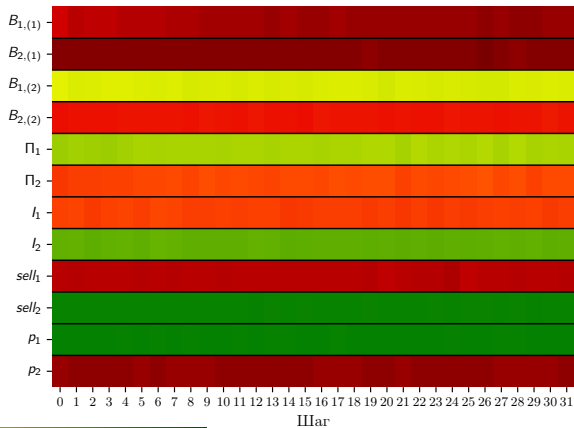


Рис.: Параметры действий фирмы 2





Эксперимент 3

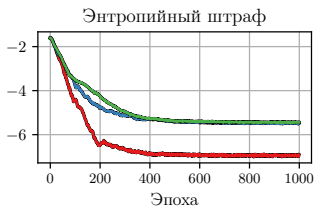
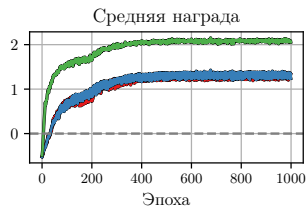
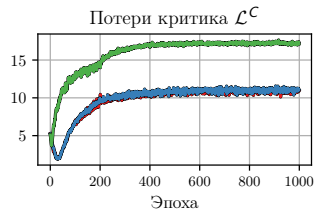
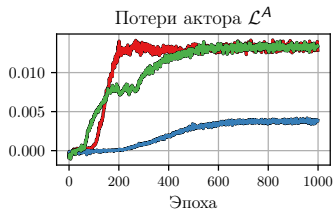
Обучение модели. «Производство и финансы, 3 отрасли»

Таблица: Функции фирм в эксперименте 3

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

Таблица: Параметры эксперимента 3

Параметр	Значение
d (срок жизни 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 данных	int64

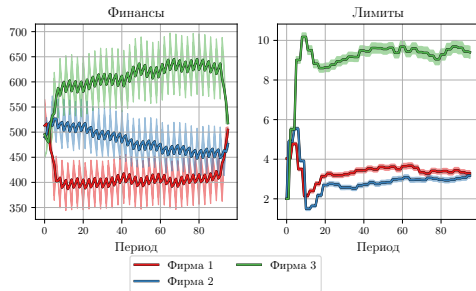
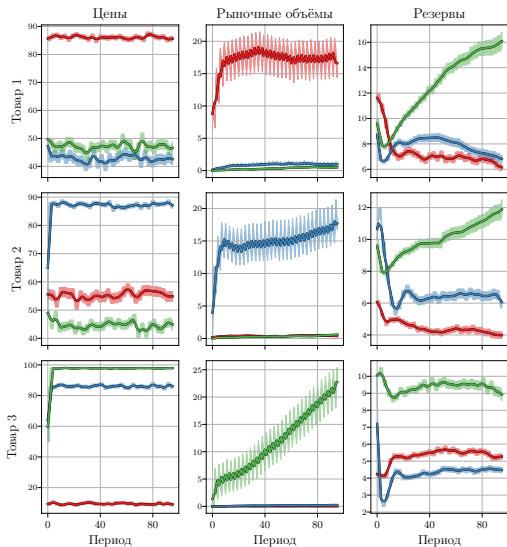


— Фирма 1 — Фирма 3
— Фирма 2



Эксперимент 3

Динамика среды. «Производство и финансы, 3 отрасли»





Эксперимент 3

Политика фирм. «Производство и финансы, 3 отрасли»

Рис.: Фирма 1

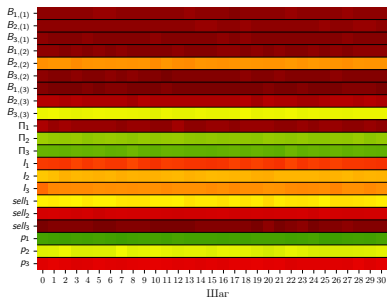


Рис.: Фирма 2

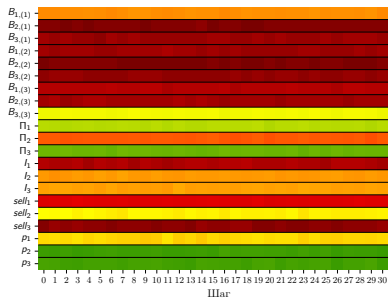
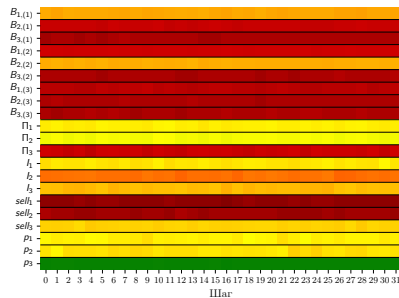


Рис.: Фирма 3 (основной капитал)





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

Параметры политик фирм:

- При усреднении, как правило, меняются несильно от эпохи к эпохе \Rightarrow агрегированные параметры действий можно описать константой или простым уравнением динамики
- Ценовая политика фирм иногда не соответствовала традиционным экономическим теориям. Фирмы могут держать высокие цены для предотвращения ошибок со стороны контрагентов.



- **Изменение архитектура 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

20 мая 2024 г.

Моделирование конкурентных стратегий гетерогенных фирм с помощью обучения с подкреплением

Выполнил: Перепелкин Владимир Андреевич БЭК205

Научный руководитель: Пильник Николай Петрович



Литература

-  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>.



Приложение. Семплирование параметров действий

Алгоритм получения параметров действий

Алгоритм. Получение параметров действий.

- 1: $a_j^{sale} \sim \text{Beta}(\alpha_{1,j}, \beta_{1,j})$ ▷ Доли резервов на продажу
 - 2: $a_j^{prices} \sim \text{Beta}(\alpha_{2,j}, \beta_{2,j})$ ▷ 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. Multi-Agent Proximal Policy Optimization

- 1: Собрать батч 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: $V_{t,i}^{clip} \leftarrow \text{clip}(V_{t,i}(s; \phi), V_{t,i}^{old} - 0.2, V_{t,i}^{old} + 0.2)$
- 8: $\mathcal{L}_i^C(\phi) \leftarrow \frac{1}{T} \sum_{t=1}^T \max\{L_1(V_{t,i}(s; \phi), \hat{V}_{t,i}), L_1(V_{t,i}^{clip}, \hat{V}_{t,i})\}$
- 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: $\phi \leftarrow \arg \min_{\phi} \mathcal{L}_i^C(\phi)$
- 16: **end for**
- 17: **end for**



Приложение. Эксперимент 4.

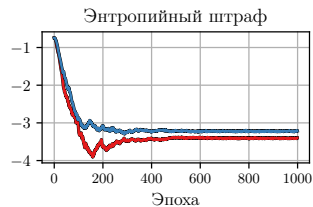
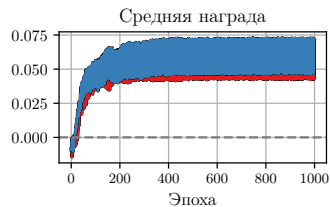
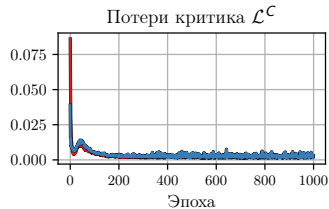
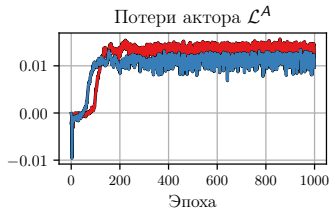
Обучение модели. «Производство и финансы, 2 асимметричные отрасли»

Таблица: Функции фирм в эксперименте 4

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

Таблица: Параметры эксперимента 4

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 данных	int64

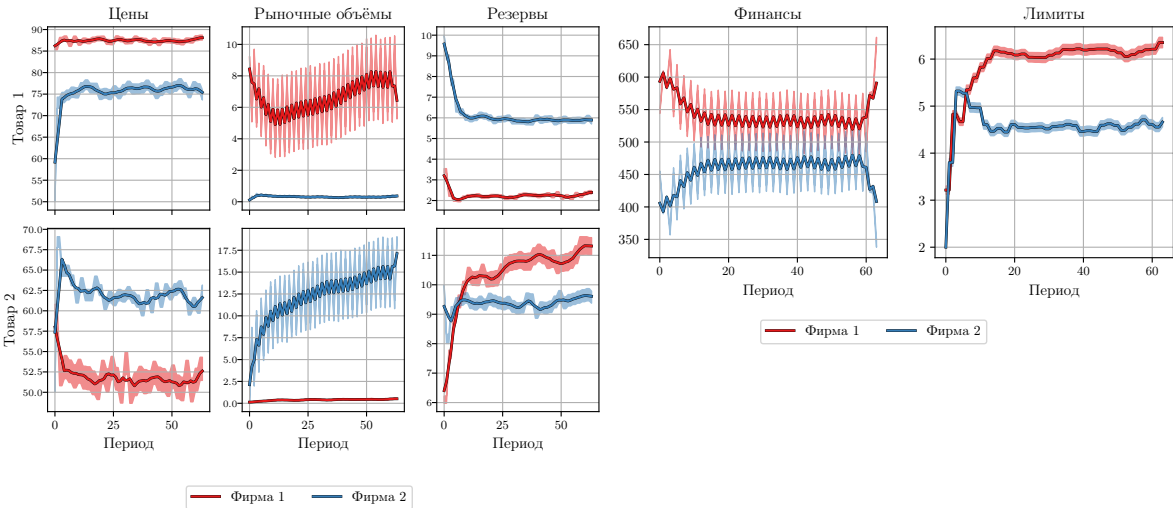


— Фирма 1 — Фирма 2



Приложение. Эксперимент 4.

Динамика среды. «Производство и финансы, 2 асимметричные отрасли»





Приложение. Эксперимент 4.

Политика фирм. «Производство и финансы, 2 асимметричные отрасли»

Рис.: Фирма 1 (дефицит)

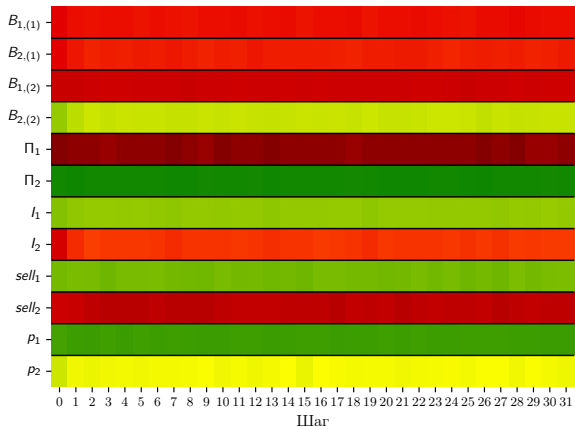
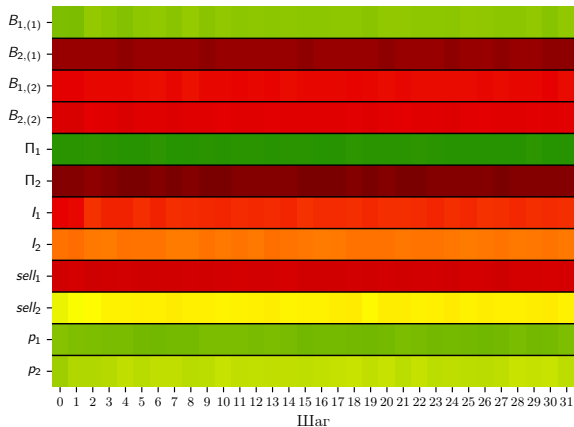


Рис.: Фирма 2





Приложение. Эксперимент 5.

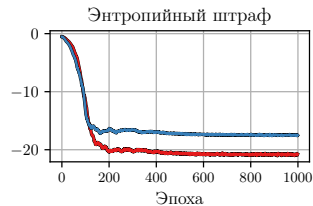
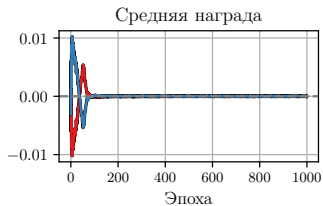
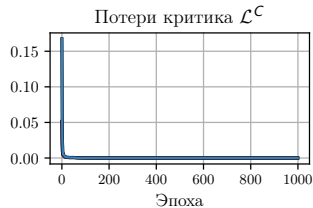
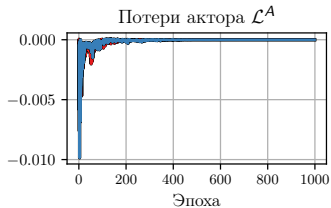
Обучение модели. «Финансы, 2 симметричные отрасли»

Таблица: Функции фирм в эксперименте 4

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

Таблица: Параметры эксперимента 4

Параметр	Значение
Основной капитал	Отсутствует
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{fin}
Тип данных	int64



— Фирма 1 — Фирма 2



Приложение. Эксперимент 5

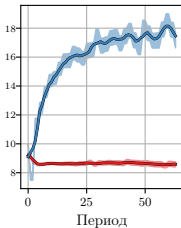
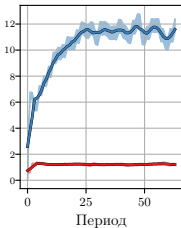
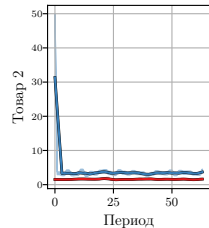
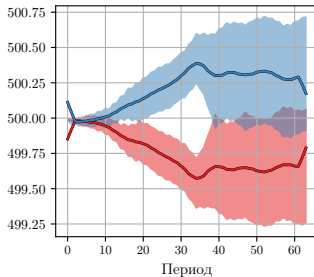
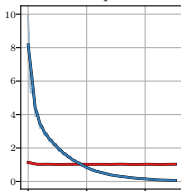
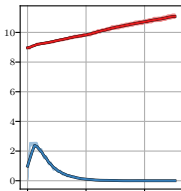
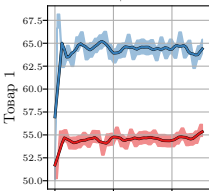
Динамика среды. «Финансы, 2 симметричные отрасли»

Цены

Рыночные объёмы

Резервы

Финансы



— Фирма 1 — Фирма 2



Приложение. Эксперимент 5.

Политика фирм. «Финансы, 2 симметричные отрасли»

Рис.: Фирма 1 (дефицит)

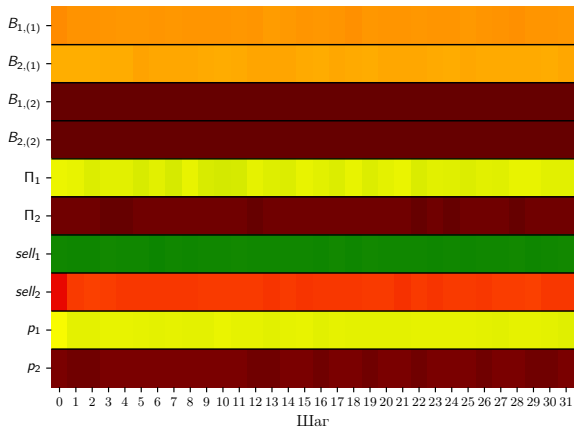
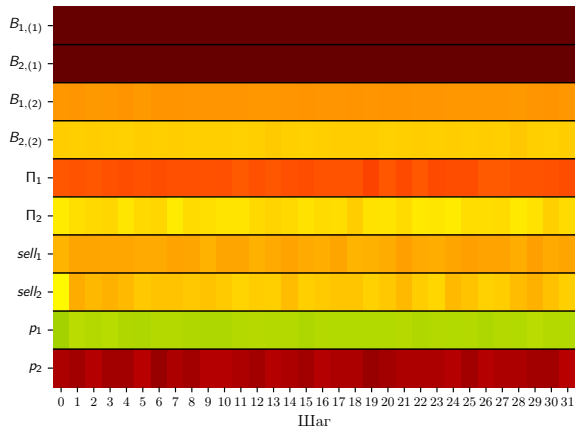


Рис.: Фирма 2





Приложение. Эксперимент 6.

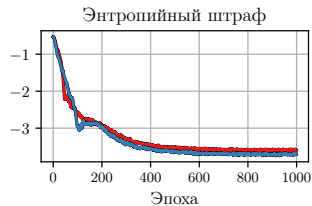
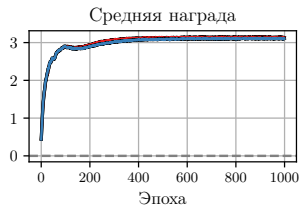
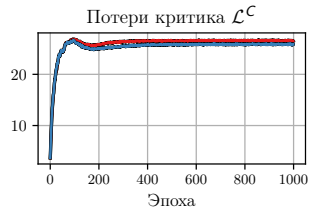
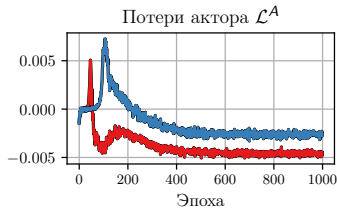
Обучение модели. «Производство, 2 симметричные отрасли»

Таблица: Функции фирм в эксперименте 6

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

Таблица: Параметры эксперимента 6

Параметр	Значение
Основной капитал	Отсутствует
Начальные резервы	(10, 10)
Начальные цены	50
Начальные финансы	500
Интервал цен	$p \in [1, 100]$
Итераций	32
Функция награды	R^{prod}
Тип данных	int64

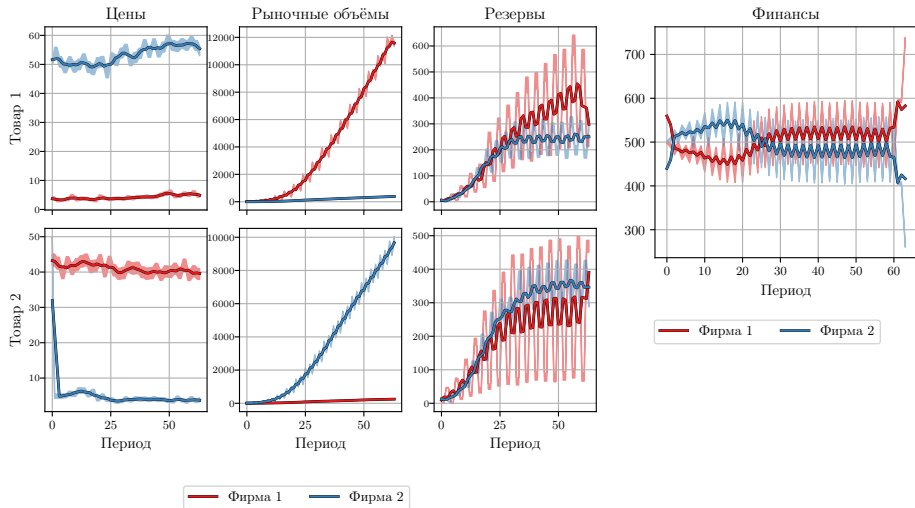


— Фирма 1 — Фирма 2



Приложение. Эксперимент 6

Динамика среды. «Производство, 2 симметричные отрасли»





Приложение. Эксперимент 6.

Политика фирм. «Производство, 2 симметричные отрасли»

Рис.: Фирма 1 (дефицит)

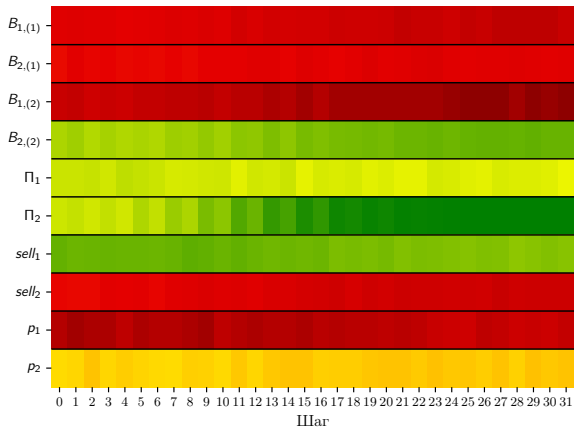
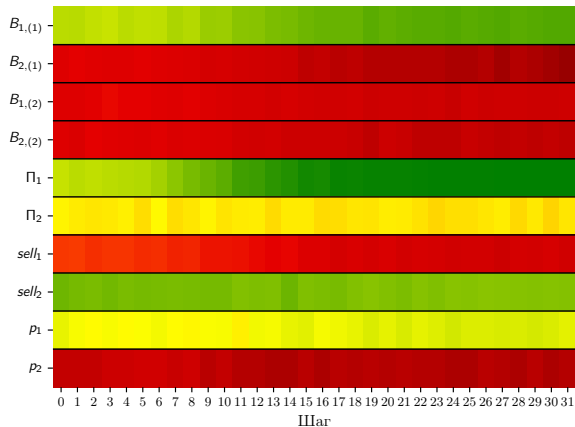


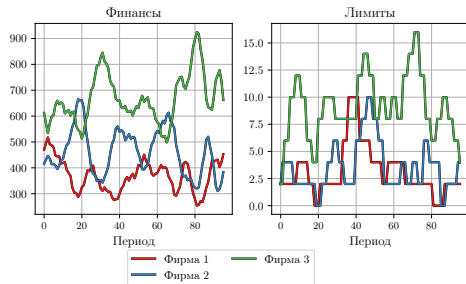
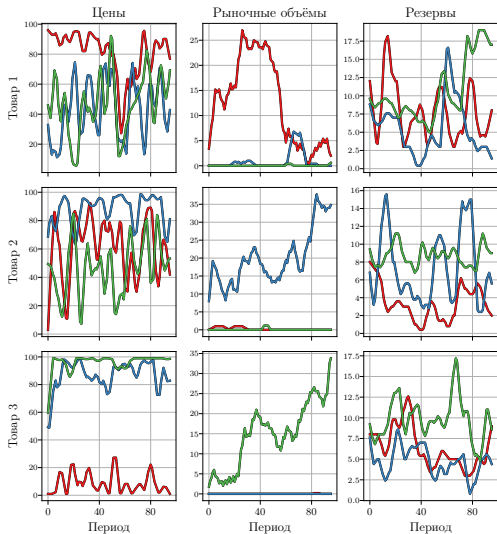
Рис.: Фирма 2





Приложение. Пример запуска без агрегирования

Динамика среды.





Приложение. Пример запуска без агрегирования

Политика фирм.

Рис.: Фирма 1

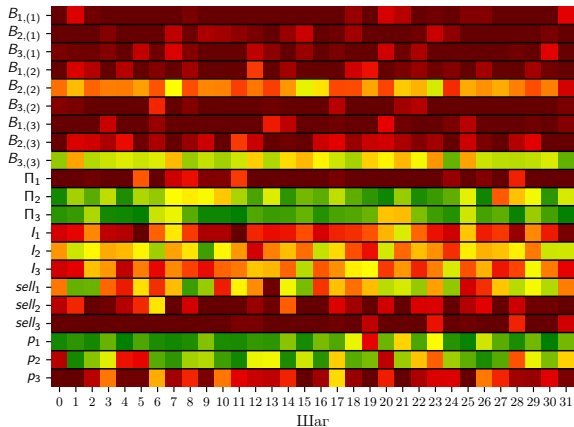


Рис.: Фирма 2

