

Розподіл ресурсів між пулами та розподіл часу між завданнями всередині пулу із застосуванням машинного навчання з підкріпленням

Мотивація

Використання методів розподілення ресурсів

- ▶ Кластери
- ▶ Хмарні обчислення
- ▶ Інше

Основна суть задачі

- ▶ змоделювати роботу планувальника ресурсів
- ▶ покращити модель

Reinforcement learning

Марковський процес прийняття рішень (S, A, r, T) , де

- ▶ S – фінітна множина станів
- ▶ A – фінітна множина дій
- ▶ $r : S \times A \times S \mapsto \mathfrak{R}$ – функція винагороди
- ▶ $T : S \times A \mapsto PD(S)$ – функція переходу станів, що відображає поточний стан та дію агента у набір ймовірнісних розподілів на S

Temporal difference(TD) learning

Функція корисності:

$$V^\pi(\mathbf{s}_t) \leftarrow V^\pi(\mathbf{s}_t) + \alpha_t(r(\mathbf{s}_t, \pi(\mathbf{s}_t), \mathbf{s}_{t+1}) + \gamma V^\pi(\mathbf{s}_{t+1}) - V^\pi(\mathbf{s}_t))$$

Стратегія:

$$\pi'(s) = \mathit{argmax} E[r(\mathbf{s}_t, a, \mathbf{s}_{t+1}) + \gamma V^\pi(\mathbf{s}_{t+1})]$$

Модель

Наближення функції цінності:

$$\hat{V}(s, p) = \sum_{i=1}^M p^i \phi^i(s)$$

де: $p_{t+1}^i = p_t^i + \alpha_t [r(s_t, \pi(s_t), s_{t+1}) + \gamma \hat{V}(s_{t+1}, p_t) - \hat{V}(\bar{s}_t, p_t)] \phi^i(\bar{s}_t)$
є параметрами навчання, що налаштовуються з часом.

$$\phi^i(s) = \frac{w^i(s)}{\sum_{i=1}^M w^i(s)}$$

$$\pi'(s) = \operatorname{argmax}[r(s_t, a, \bar{s}_t) + V^\pi(\bar{s}_t)]$$

Функція виграшу:

$$r(s_t, s_{t+1}) = B(s_t, s_{t+1}) + TC(s_t, s_{t+1})$$

Моделювання

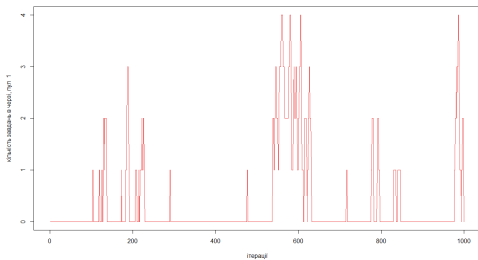
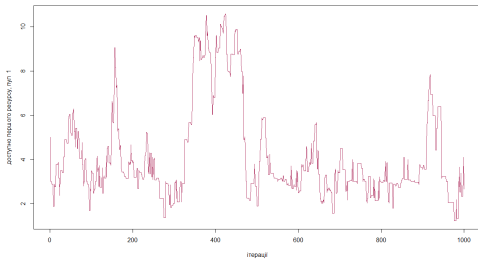
Параметри

Було змодельовано поведінку пулів при 100000 ітерацій с такими початковими значеннями:

1. початкові рейти приходу нових завдань 0.24, 0.18, 0.19
2. початковий розподіл ресурсів (6,7) для всіх трьох пулів
3. максимальний та мінімальний рейти приходу нових завдань - 0.1 та 0.3 відповідно
4. максимальний та мінімальний час виконання завдань - 7 і 10 відповідно

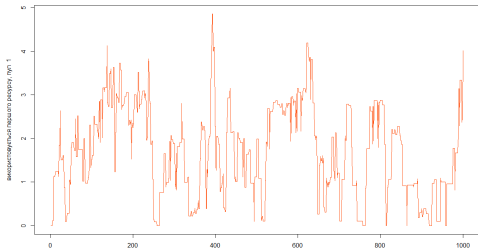
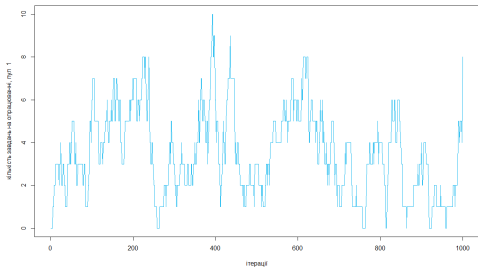
Моделювання

Результат: зміна черги та доступного ресурсу



Моделювання

Результат: зміна кількості завдань на опрацюванні та ресурсу, що використовується



Покращення

витіснення

Тип підходу	Середній час знаходження в черзі	Максимальний час знаходження в черзі
Проста з черга	4.12	140.52
Черга з витісненням	1.27	49.24

Покращення

випадковий вибір

Тип підходу	Середній час знаходження в черзі	Максимальний час знаходження в черзі
Витіснення з чергою	0.87	43.51
Витіснення з випадковим вибором завдань	0.88	44.19

Недоліки підходу

- ▶ статичне завдання параметрів
- ▶ обмеження на діапазон параметрів
- ▶ гонка між пулами