

HRL

# Проблемы RL и подходы к решению

## Hierarchical decomposition

- Большие размерности данных
- Большие объемы данных
- Большое время обучения
- Малое влияние “логики мира” на агента
- Long term credit assignment

## FeUdal Networks

- Long term credit assignment
- Sparse rewards
- Большие размерности данных

# Q-value => HQ-value iteration

Subtasks :  $O = \{O_0, O_1, \dots, O_n\}$

Subpolicy :  $\pi = \{\pi_0, \pi_1, \dots, \pi_n\}$ .

$$Q_i(s, u) = V(s, u) + \sum_{s', N} P_i^\pi(s', N | s, u) \gamma^N Q_i^\pi(s', \pi_i(s')) \quad (1)$$

$$V(s, u) = \begin{cases} \max_{u'} (Q_u(s, u')) & u \text{ is subtask} \\ \sum_{s'} P(s' | s, u) R(s' | s, u) & u \text{ is primitive} \end{cases} \quad (2)$$

# Direct acyclic graph

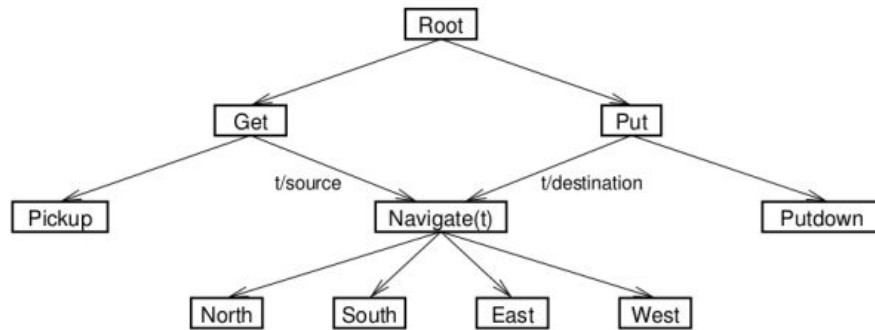


Figure 2: DAG 1

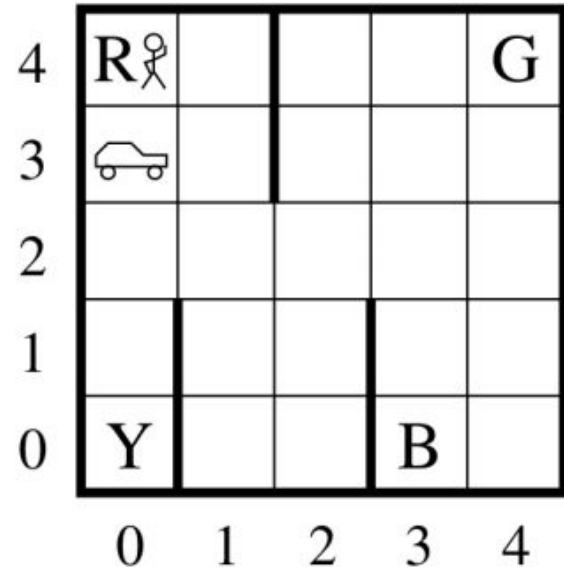


Figure 1: The Taxi Domain

# Как обучать?

- Существует топ-сорт DAG
- Учим в порядке от листьев к корню

---

**Algorithm 1** Hierarchical Q-value iteration (HQI)

---

**Require:**  $O, D$

$Train \leftarrow O_i \in O$  with only primitive children

$Done \leftarrow \{A\}$

**while**  $Train \neq empty$  **do**

**for**  $O_i \in Train$  **do**

$SQI(O_i, D)$

$Done.add(O_i)$

**end for**

$Train \leftarrow O_i \in (O - Done) \text{ AND } U_i \in Done$

**end while**

# State abstraction

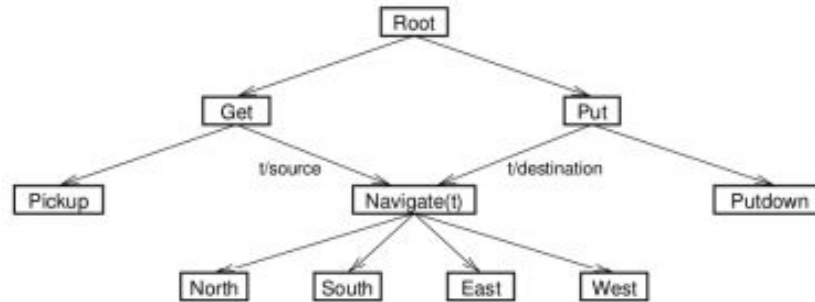
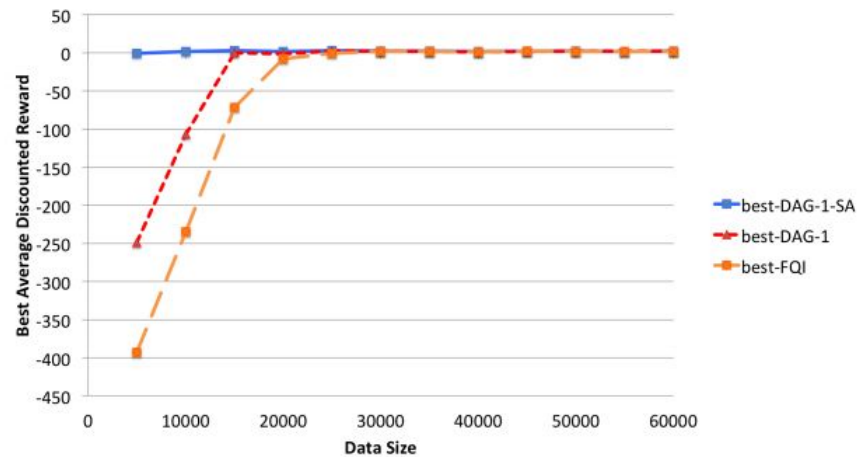
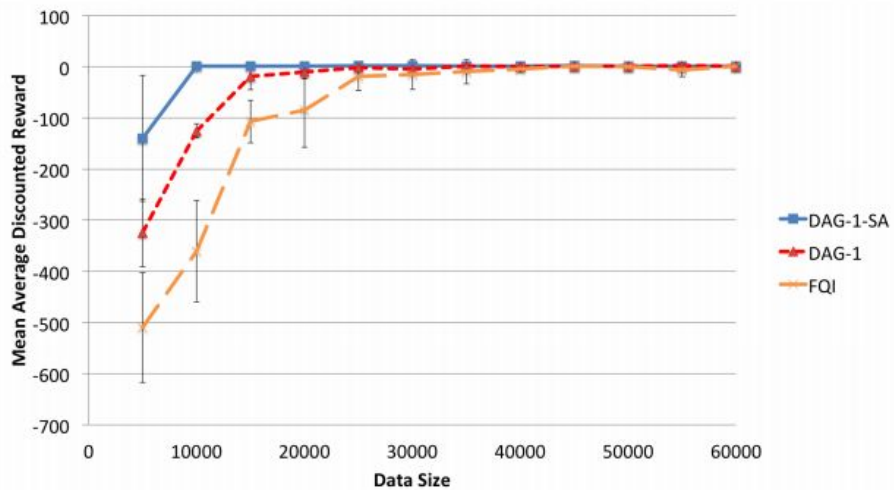


Figure 2: DAG 1

Table 1: DAG-1 State Abstraction

subtask	active states
root	[pass]
get	[pass x y]
put	[dest x y]
navi_get	[pass x y]
navi_put	[dest x y]

# HQI vs FQI vs HQI-SA



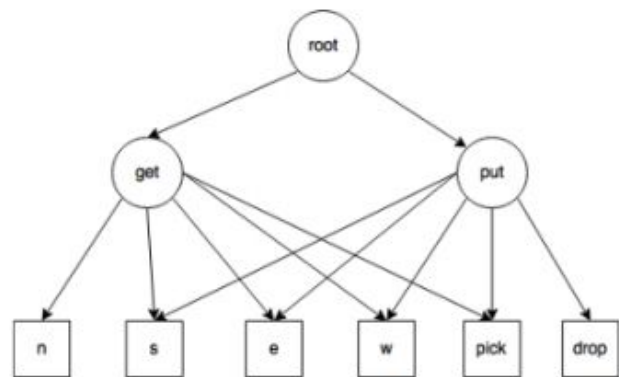


Figure 5: DAG 2

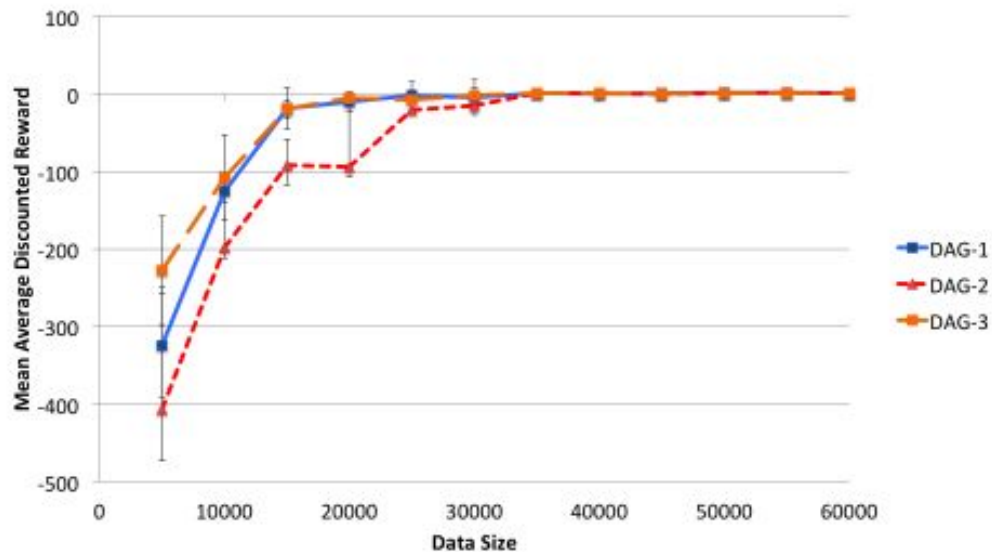
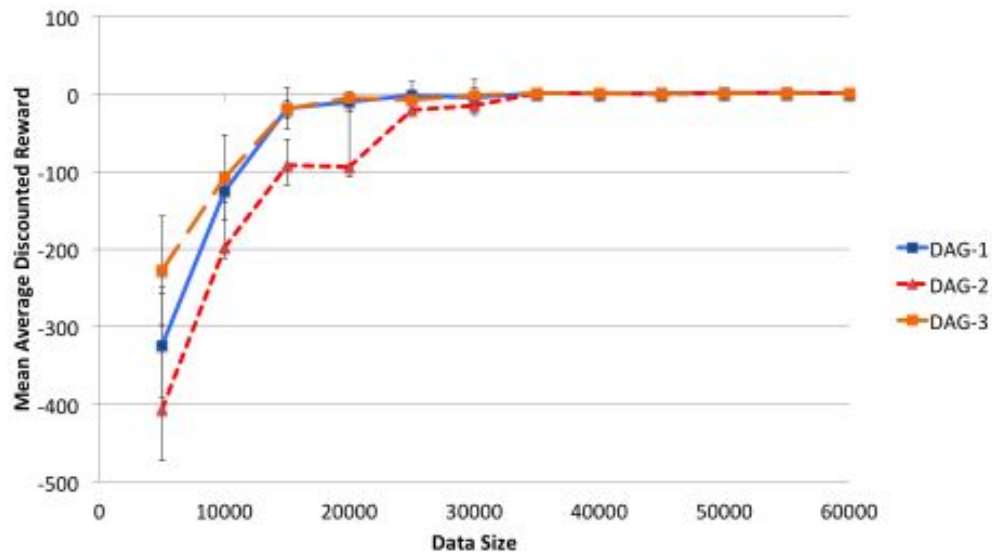
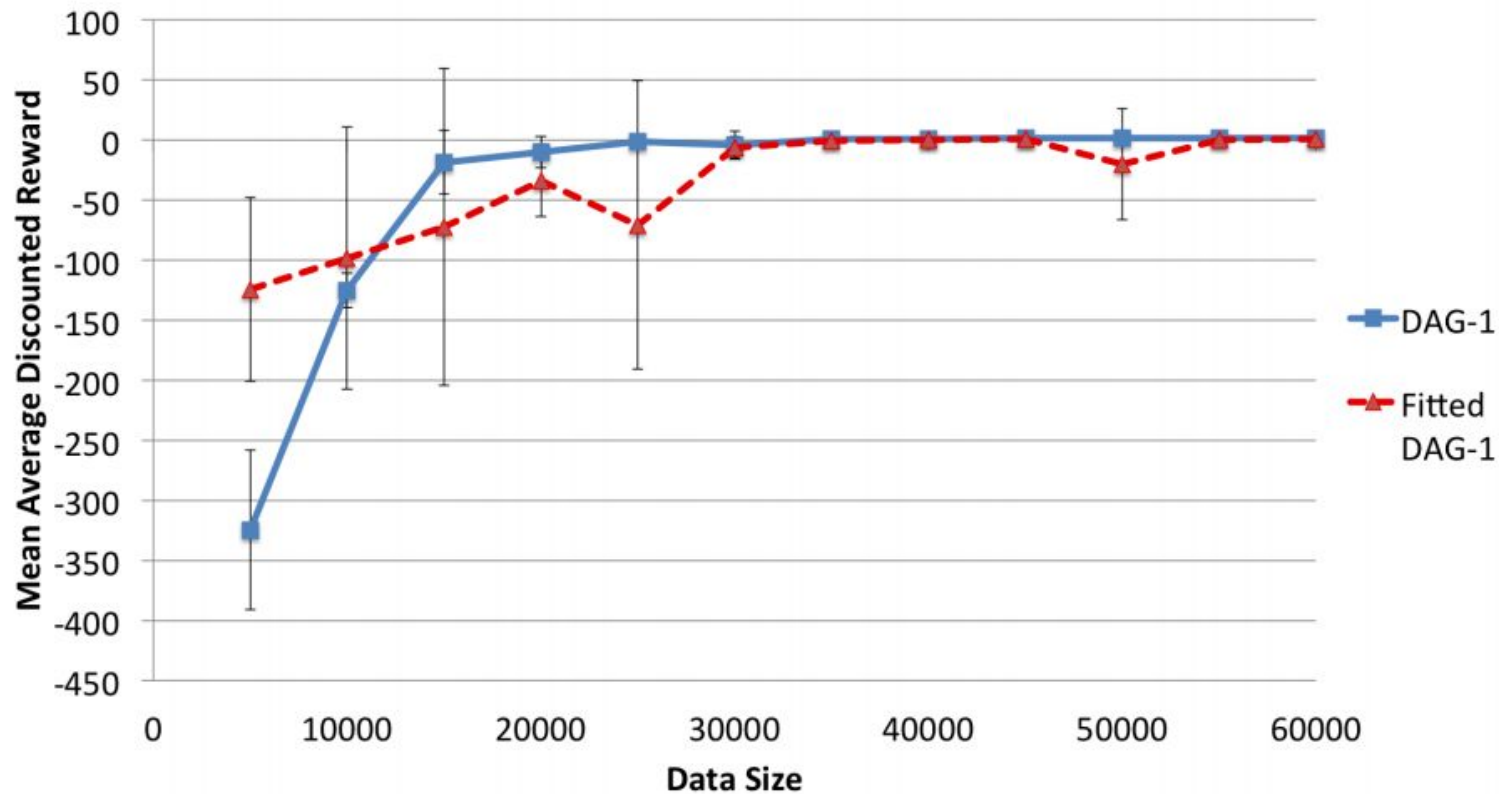


Figure 6: DAG 3

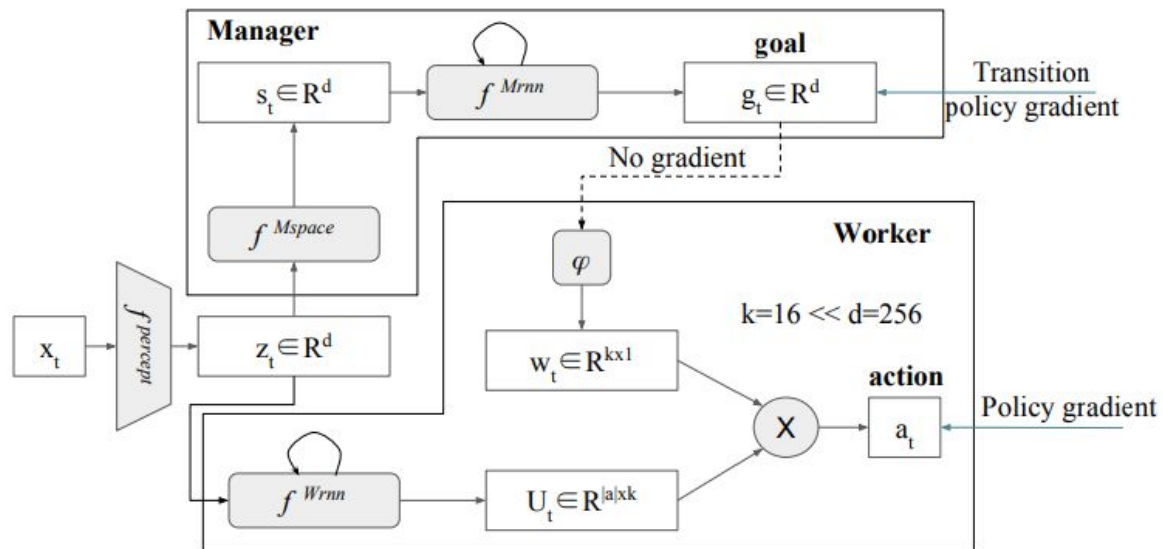




# Деревья?



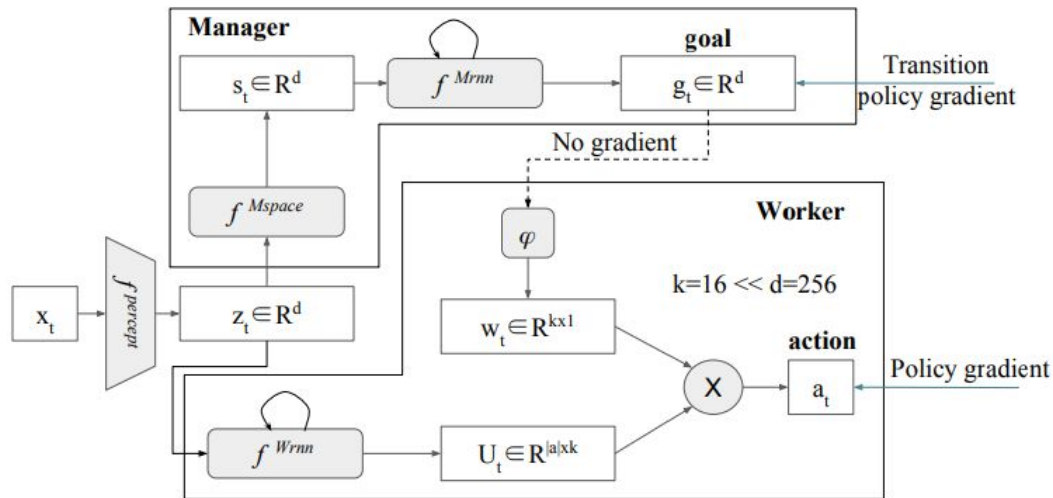
# FeUdal Networks



# Архитектура

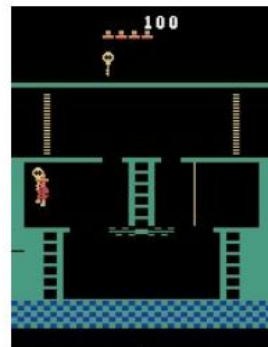
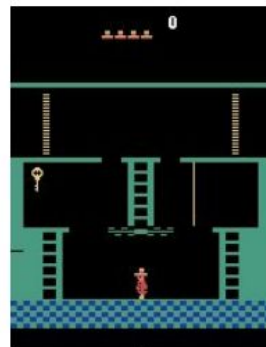
## Manager - Создаёт goals (low-dimensional)

Worker - Пытается goals выполнить





start



Goal count

0

Time step

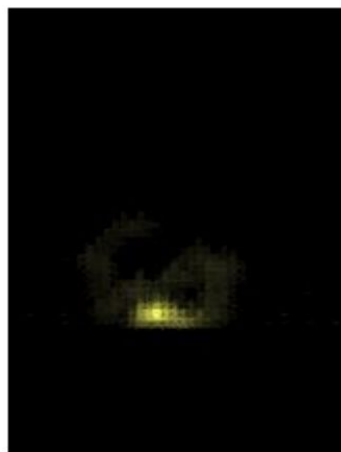
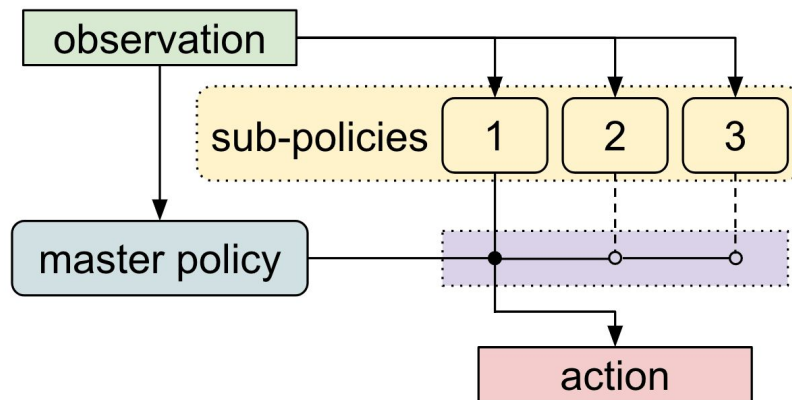
180

15

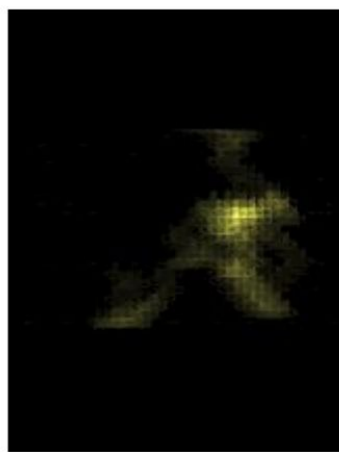
10

5

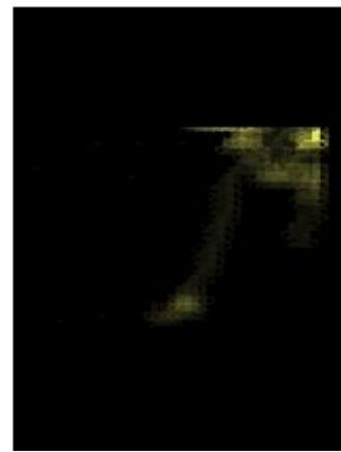
# Комбинирование



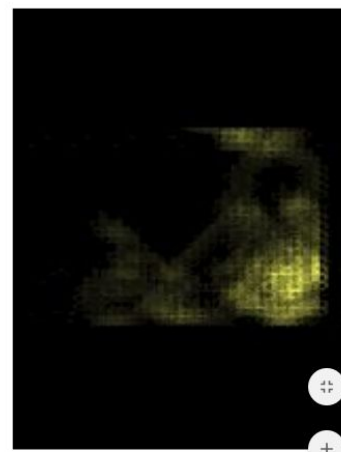
sub-policy 1



sub-policy 2



sub-policy 3



sub-policy 4

# Список литературы

<https://blog.openai.com/learning-a-hierarchy/>

<https://arxiv.org/pdf/1703.01161.pdf>

<https://arxiv.org/pdf/1603.08869.pdf>