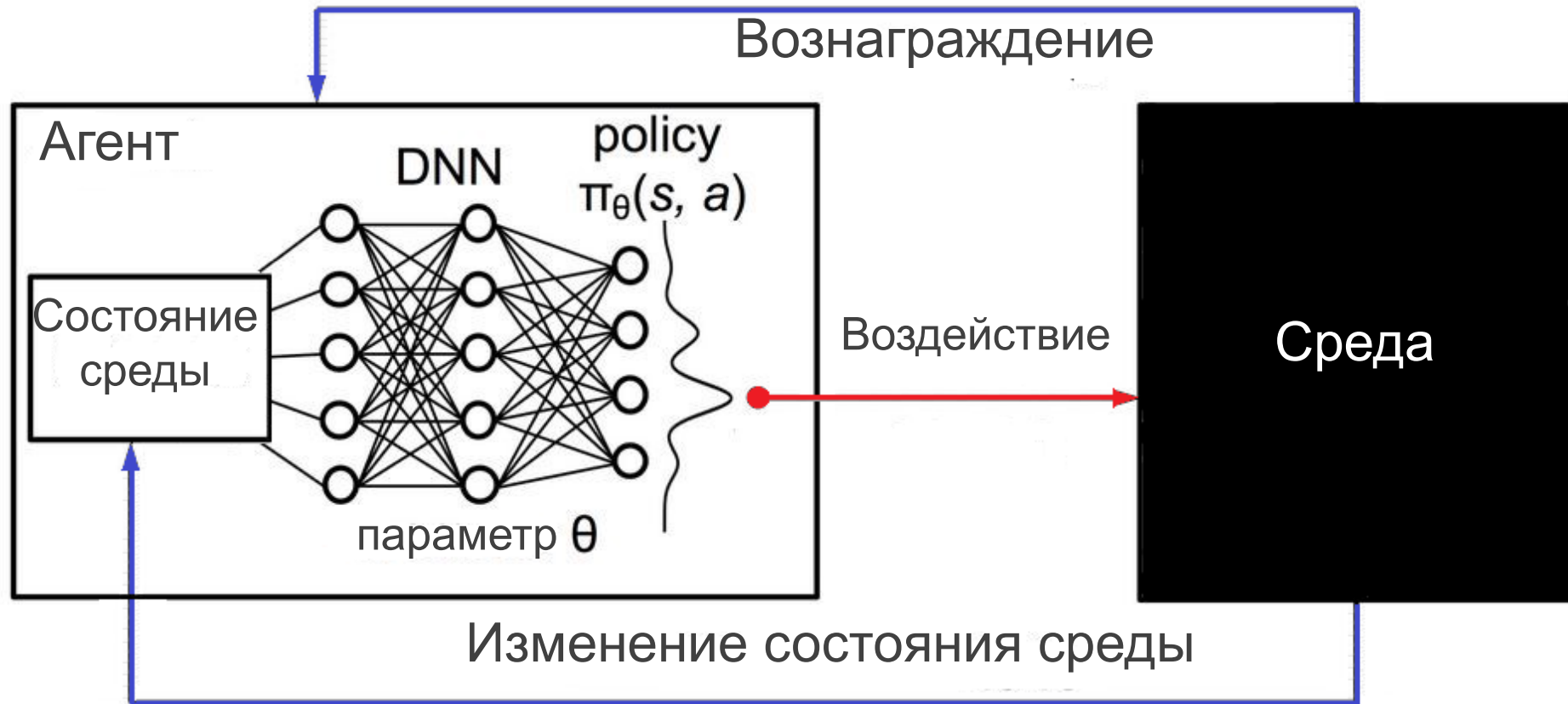


**Reinforcement Learning агент:  
управление технологическими  
процессами**

All-Over-IP  
2019

# RL агент: общая схема



Состояния среды:  $S = \{s_t\}$ , возможные действия агента  $A = \{a_t\}$

# Reinforcement Learning Агент

## Определение **Reinforcement Learning**

Политика RL агента  $\pi(\mathbf{S}_t)$  определяет оптимальное действие агента в состоянии  $\mathbf{S}_t$  для максимизации ожидаемого совокупного вознаграждения.

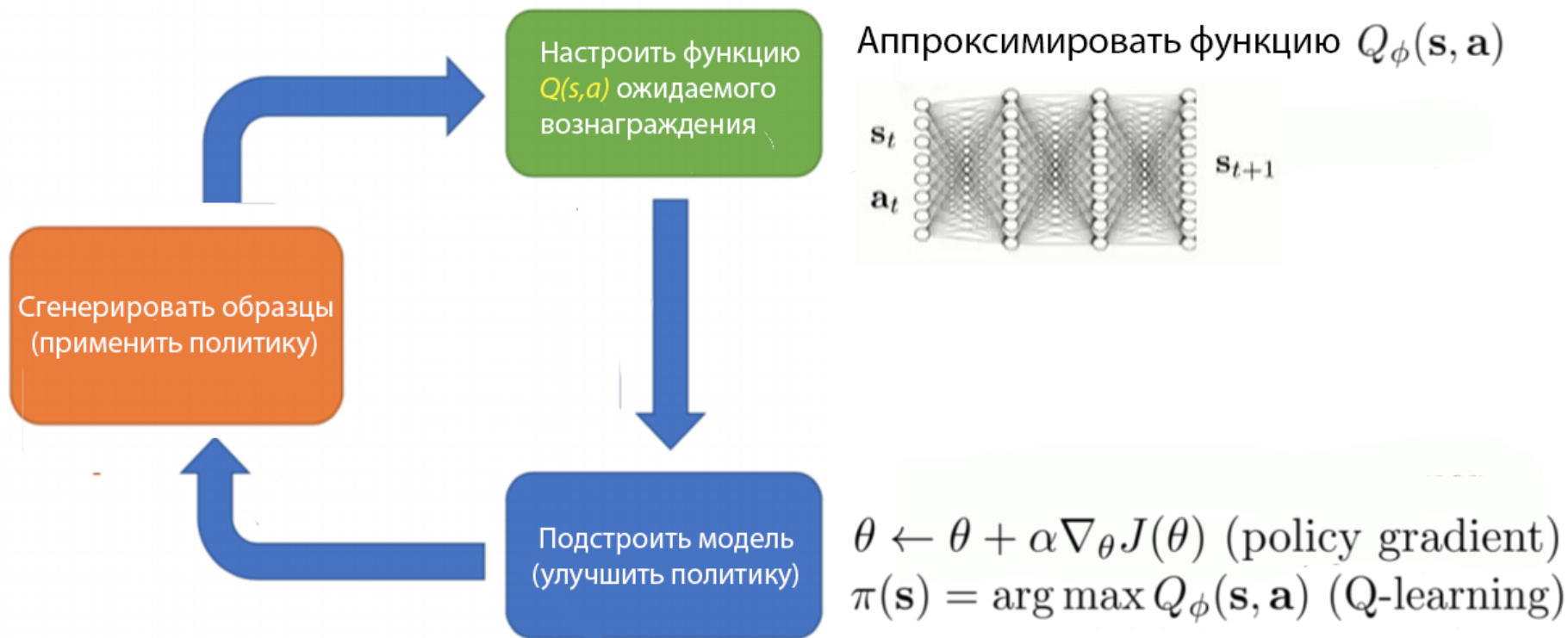
Функция стоимости состояния  $V_\pi(\mathbf{S}_t)$  предсказывает размер ожидаемого совокупного вознаграждения для состояния  $\mathbf{S}_t$  при условии, что агент будет следовать  $\pi(\mathbf{S}_t)$ .

Функция  $Q_\pi(\mathbf{S}_t, \mathbf{a}_t)$  также предсказывает размер ожидаемого совокупного вознаграждения для состояния  $\mathbf{S}_t$  при условии, что агент выберет действие  $\mathbf{a}_t$  и будет следовать политике  $\pi(\mathbf{S}_t)$ .

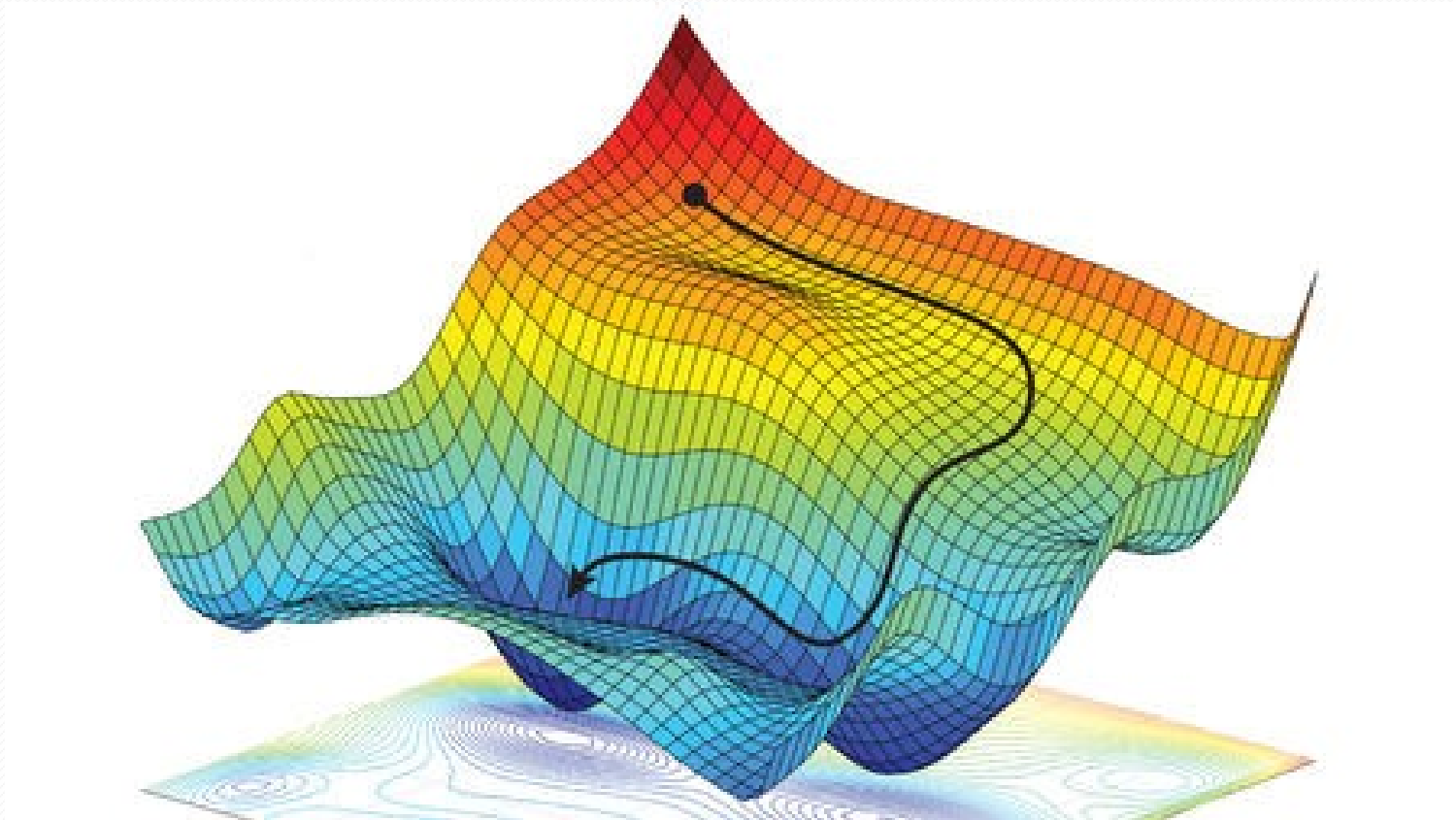
## Цель **Reinforcement Learning**

RL Агент должен самообучиться одновременно и  $\pi(\mathbf{S}_t)$ , и  $V_\pi(\mathbf{S}_t)$

# Постановка задачи RL агента



Что есть градиент в случае RL агента?



# Пример: управление отоплением дома

Постановка задачи:

**Агент:** *Python, Keras/TensorFlow, PG RL agent*

**Объект(среда):** многоэтажный жилой дом<sup>1</sup>

**Воздействия:** управление задвижкой подачи горячей воды

**Модель среды(агент):** 10 последних значений температуры на выходе<sup>2</sup>

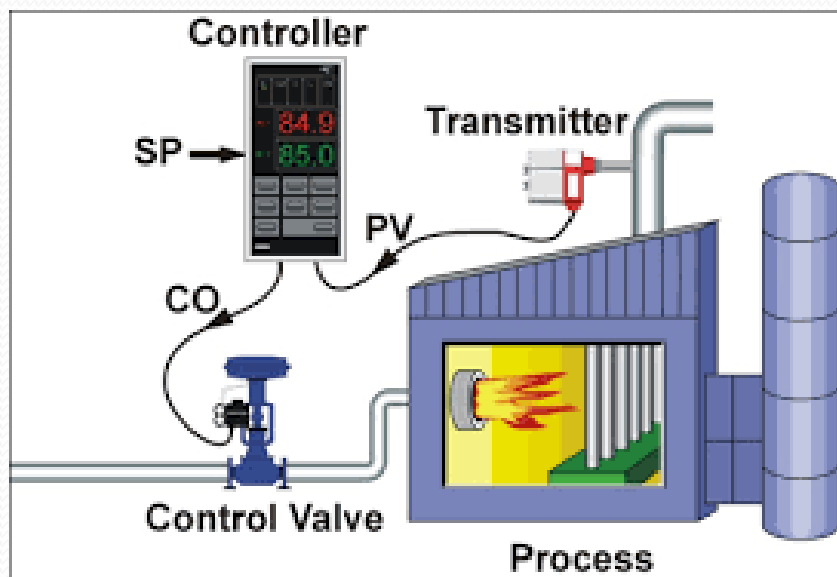
**Вознаграждение:** штраф за расхождение с заданным графиком

<sup>1</sup> В версии 1.0 объект представлен симулятором (на языке Python, OpenAI environment)

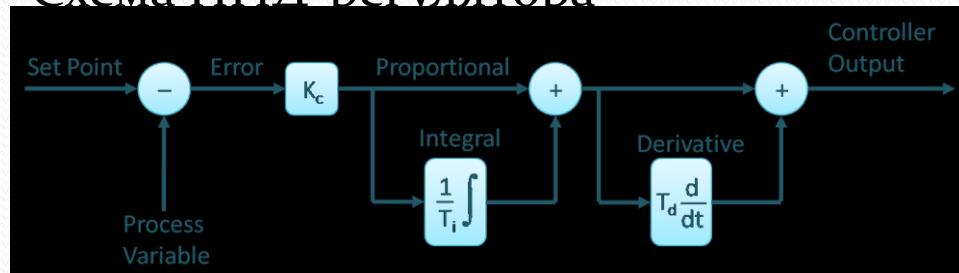
<sup>2</sup> Агент версии 2.0 будет иметь в модели среды прогноз погоды на сегодня

# Пример: управление отоплением дома

Сегодня отоплением дома управляет автомат – ПИД регулятор



## Схема ПИД-регулятора

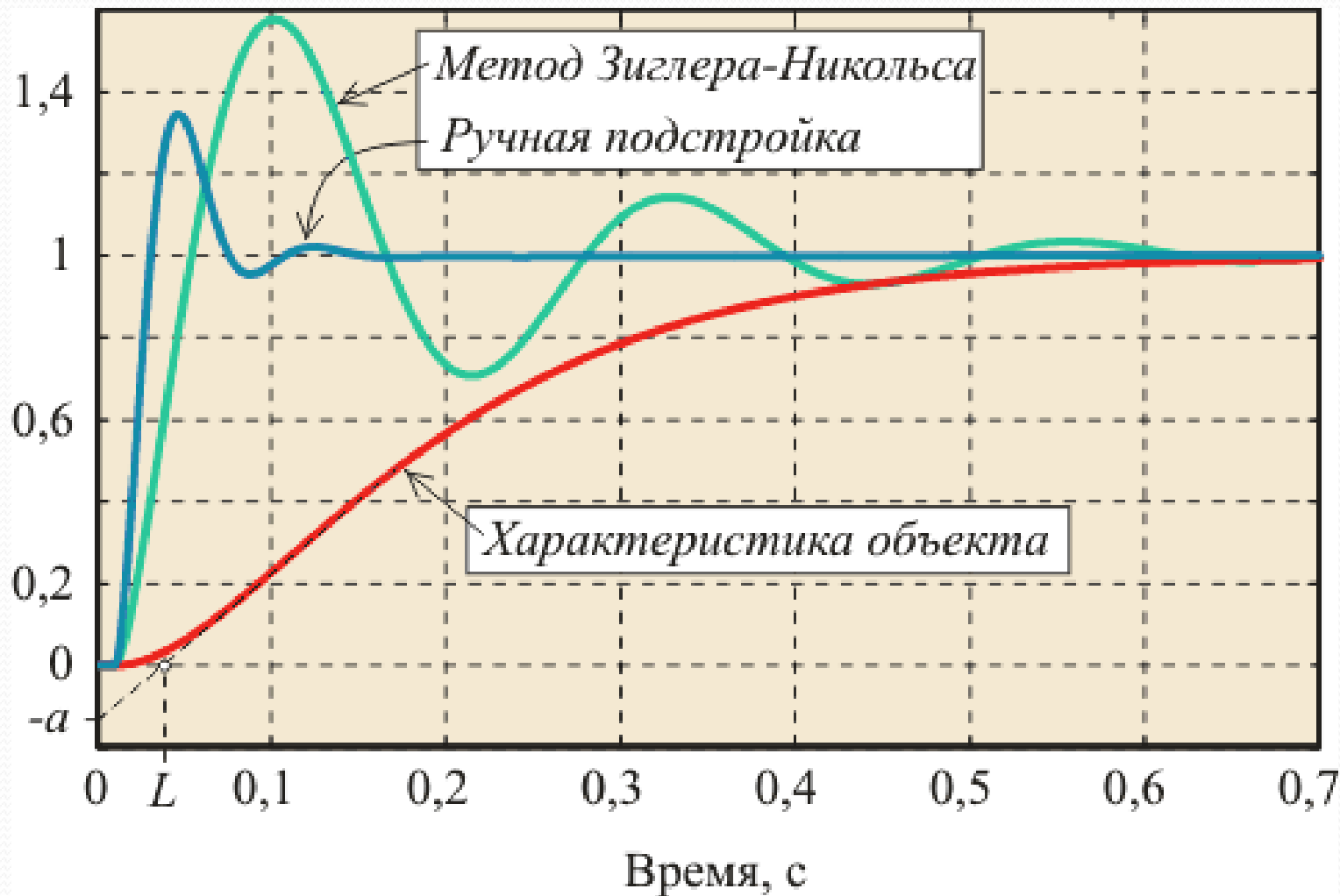


Управляющий сигнал ПИД регулятора это сумма трёх слагаемых:

- 1) П — пропорционально сигнала рассогласования
- 2) И — интеграл сигнала рассогласования
- 3) Д — производная сигнала рассогласования

# Результат управления ПИД-регулятора:

Единичная «ступенька» - поднять температуру на 1 единицу





# Математический подход

Модель управления ПИД регулятором на базе классической математики требует решения системы дифференциальных уравнений **четвертого порядка**

**PID Control:** (pure PID)

$$G_C(s) = K_C \left( 1 + \frac{1}{\tau_I s} + \tau_D s \right)$$

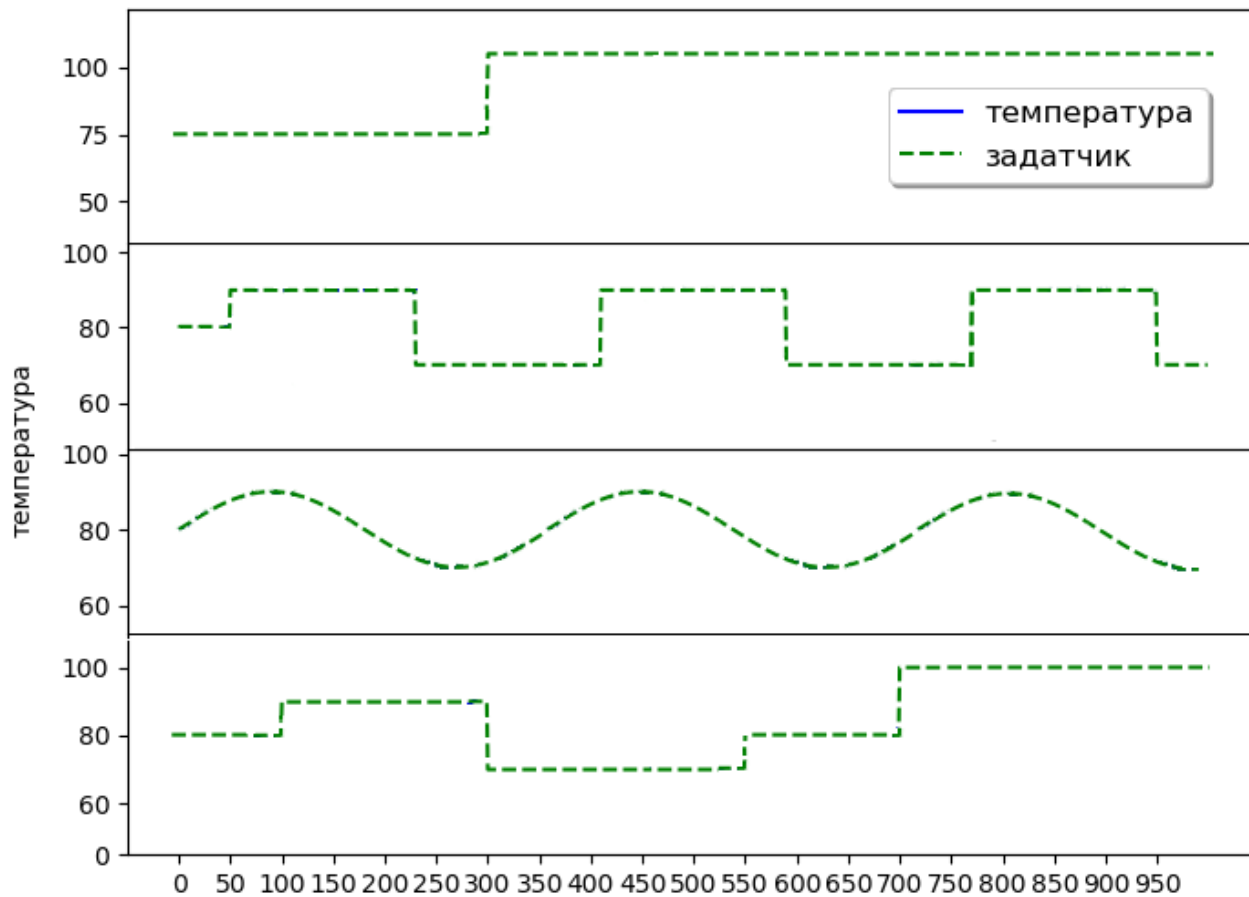
No offset, adjust  $K_C$ ,  $\tau_I$ ,  $\tau_D$  to obtain satisfactory result (requires solving for roots of **4<sup>th</sup> order** characteristic equation).

∴ Analysis of roots of characteristic equation is one way to analyze controller behavior

$$1 + G_C G_V G_P G_M = 0$$

# Управление отоплением дома

## Набор стандартных тестов\* для ПИД регулятора



Ступенька

Меандр

Синус

Нерегулярный меандр

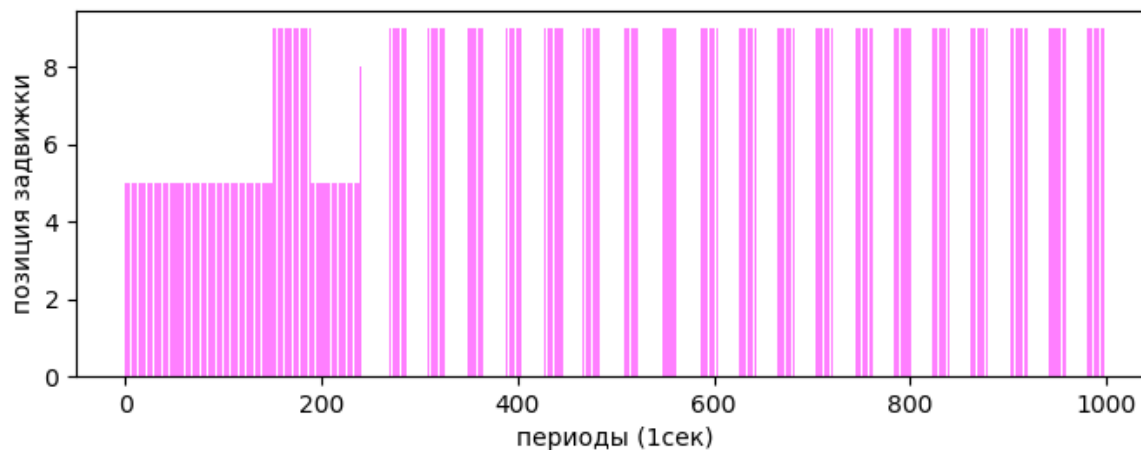
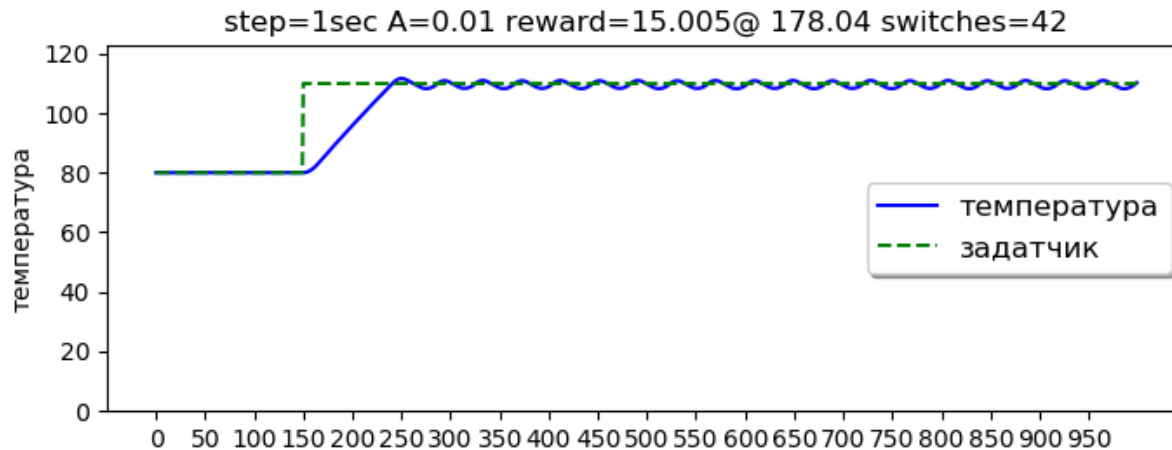
\* *Задается плановый график температуры управляемого объекта (задатчик) который должен выдержать ПИД регулятор*

# Умное управление отоплением дома

ИИ метод – PG агент (Gradient Policy), unsupervised machine learning

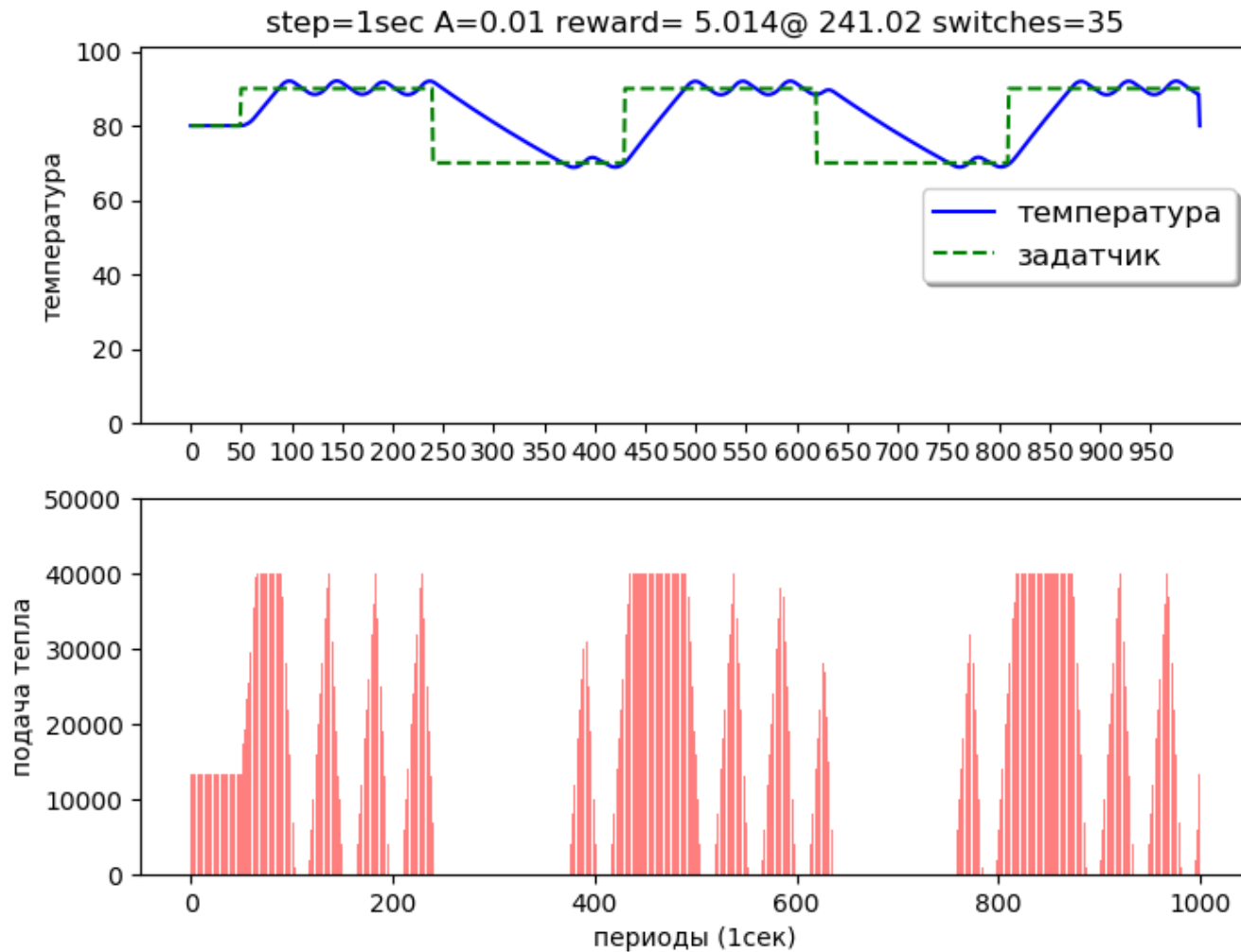
Реализован на языке Python 3.6, ИИ средства - Keras, TensorFlow

## Ступенька



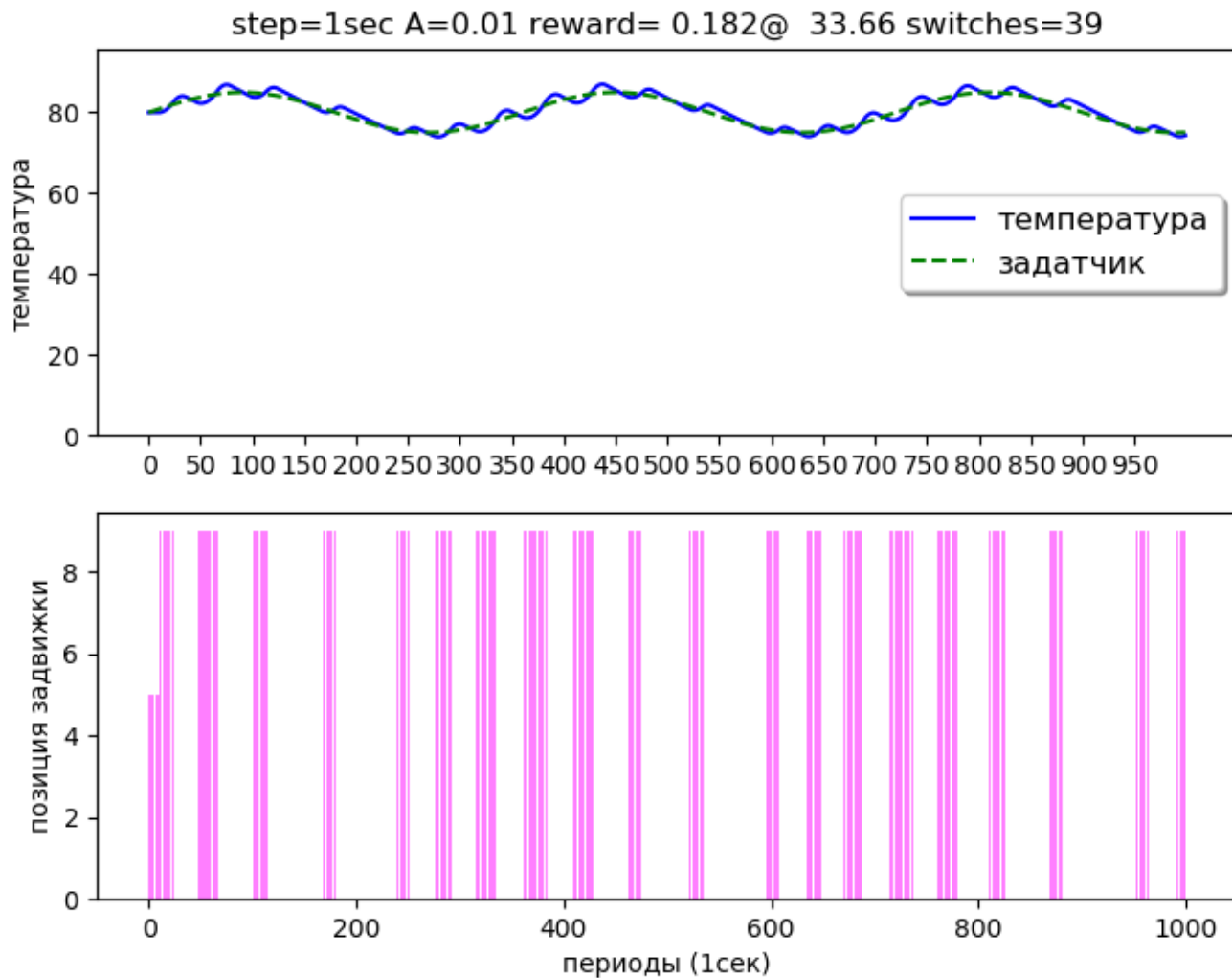
# Умное управление отоплением дома

## Меандр

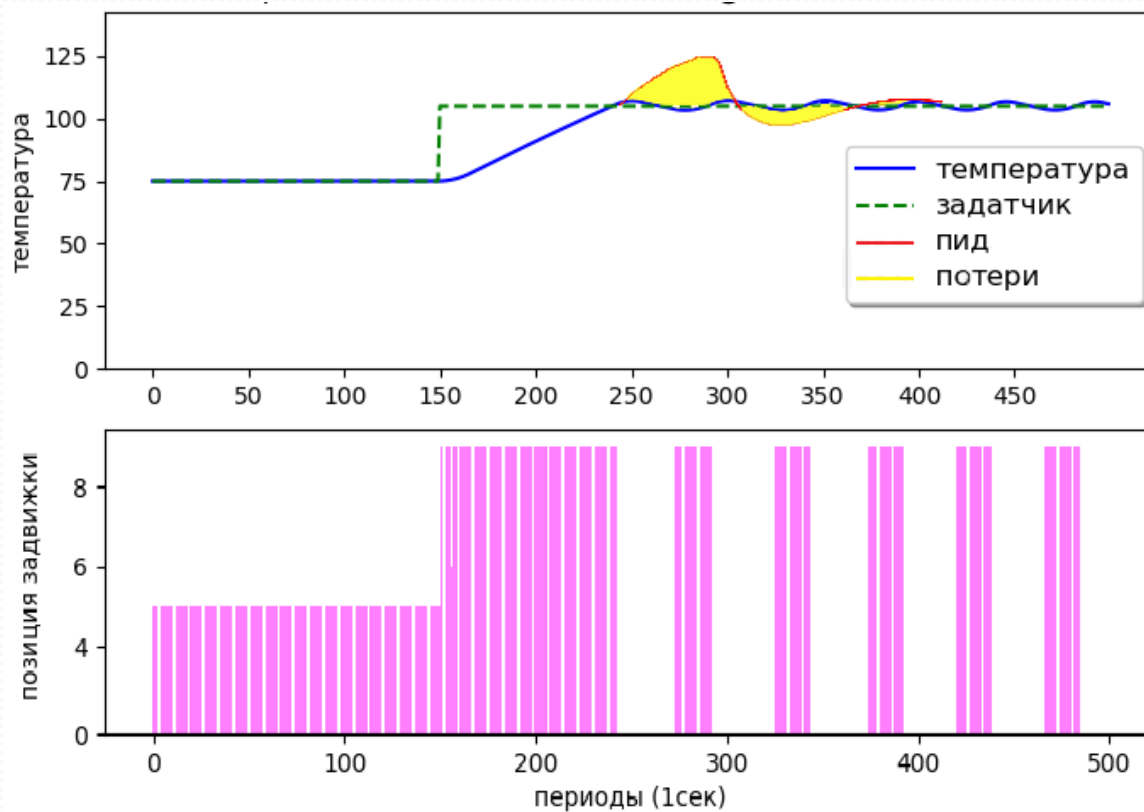


# Умное управление отоплением дома

## Синус

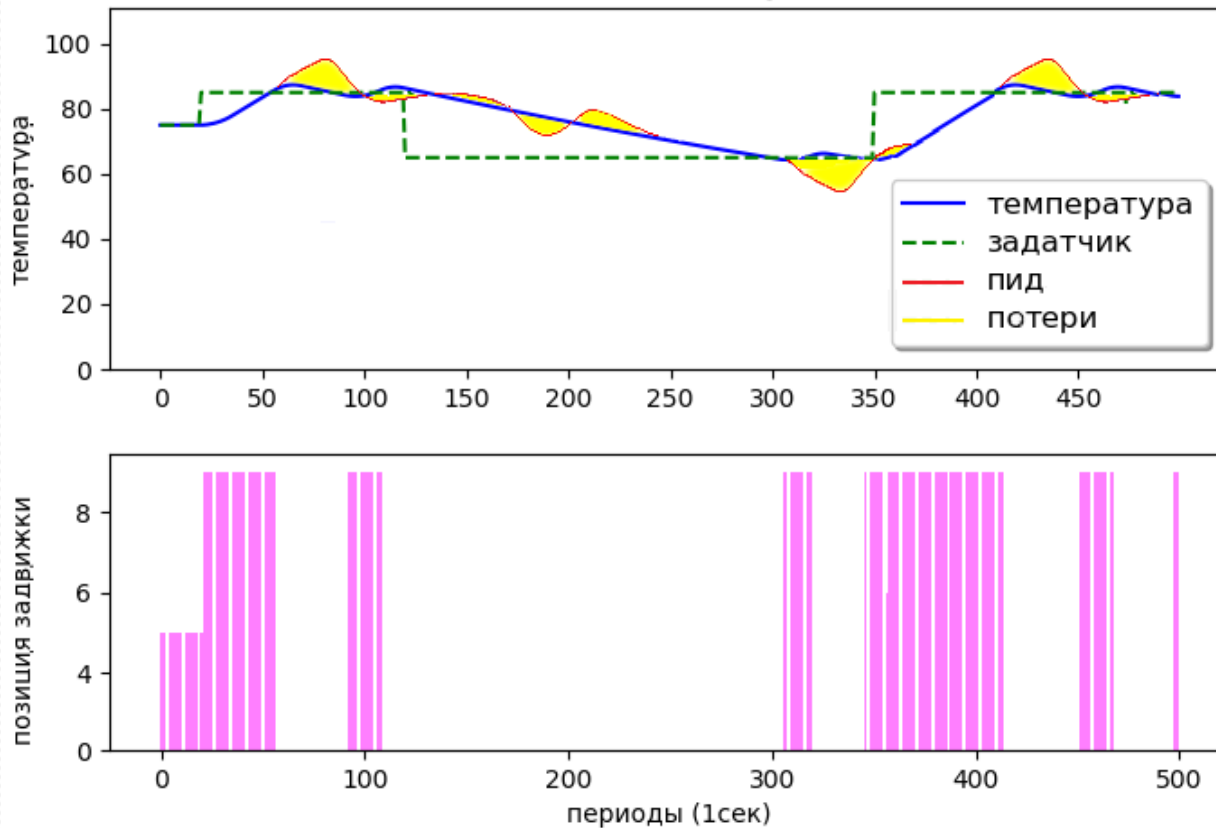


# Смысл применения RL агента?



- 1) ПИД-регулятор «проскакивает» выше задатчика - перегрев и потеря времени
- 2) ПИД-регулятор «проскакивает» ниже задатчика – переохлаждение, дополнительные расходы тепла и потеря времени

# Смысл применения RL агента?

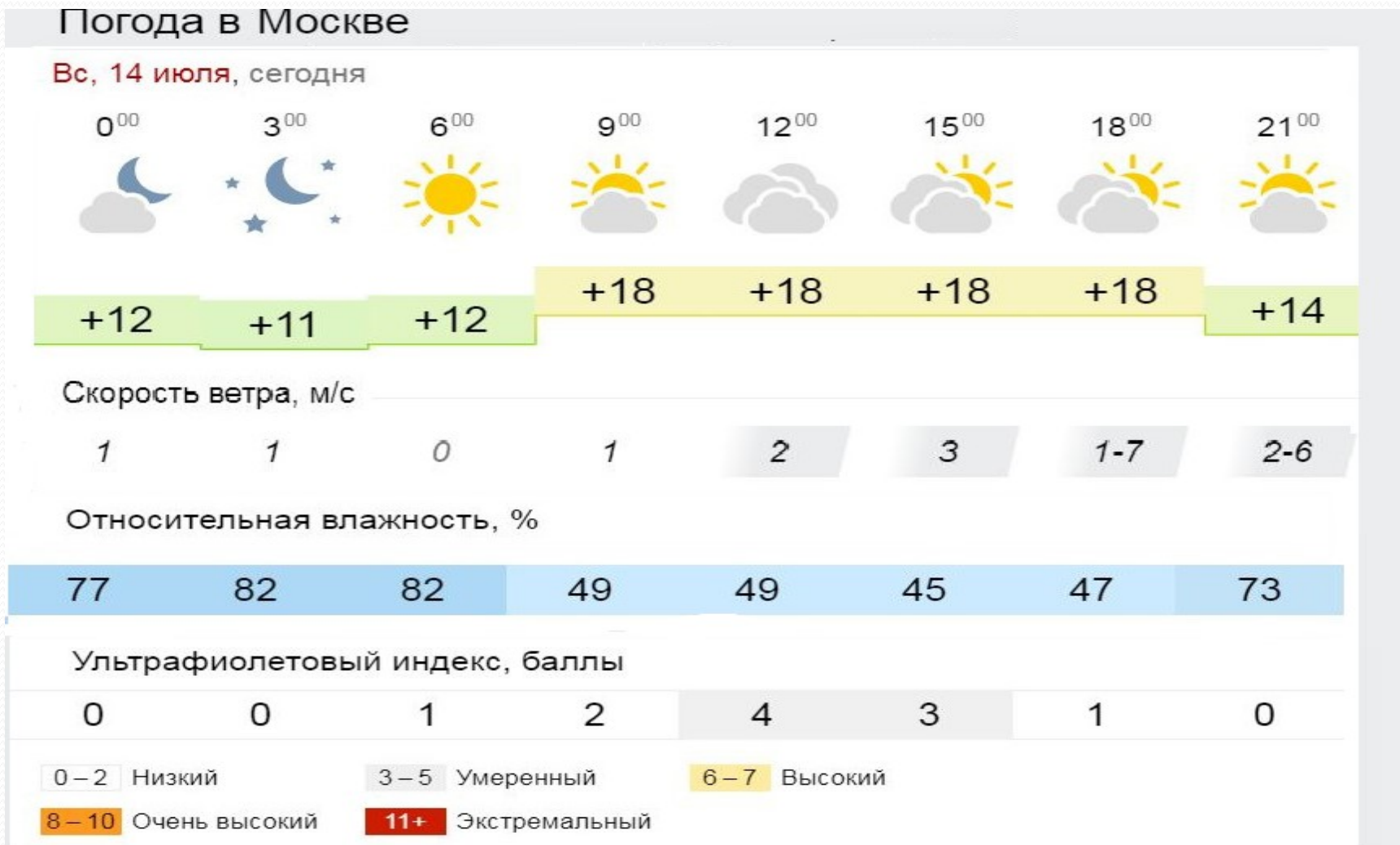


Если задатчик меняется с высокой частотой, то ПИД регулятор может «потеряться». Высокие частоты можно убрать фильтром, но тогда ПИД регулятор будет игнорировать задатчик и не выполнит свою функцию.

Интеллектуальный агент выдерживает бОльшие частоты задатчика

# RL агент - версия 2.0:

В модель среды Агент 2.0 будет включен прогноз погоды из «облака»

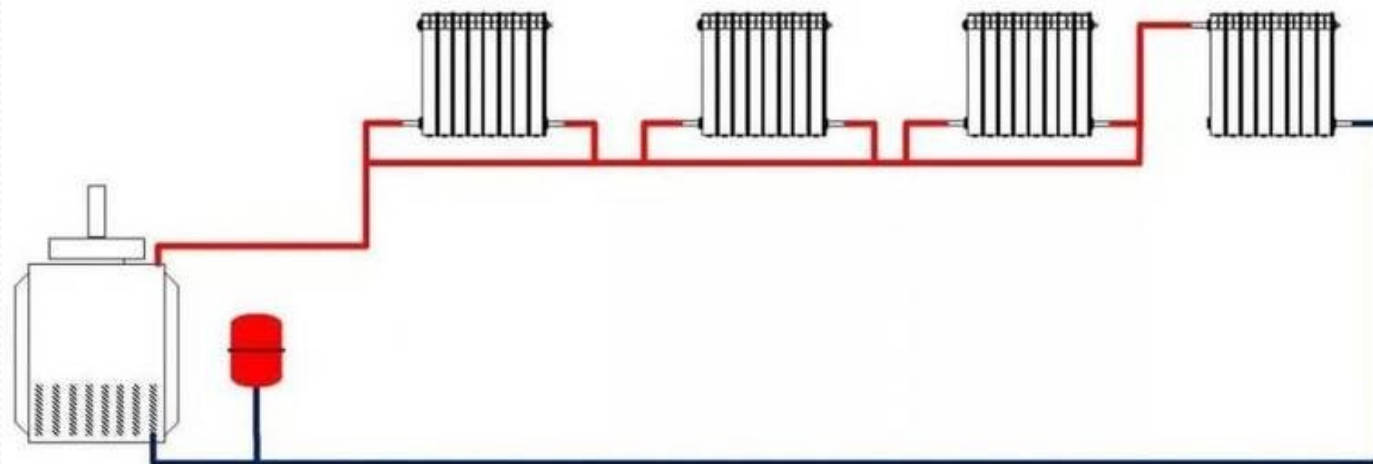




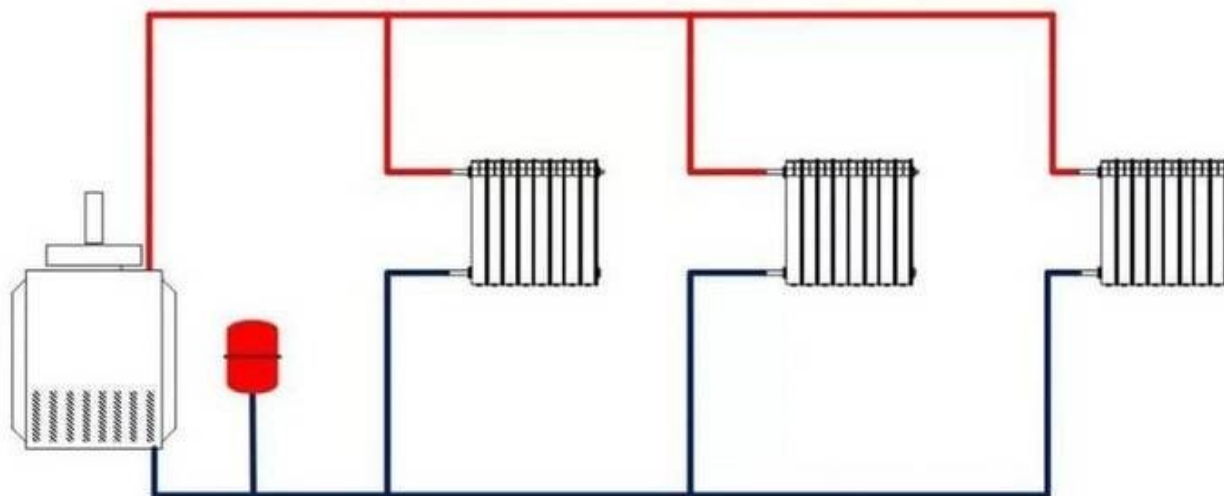
# ИИ в ЖКХ: примеры применения

Кто	Что	ИИ метод	Эффект
National University of Singapore, Singapore, 2017	Тепло	Deep RL, gradient policy	15%
University of California, Riverside, CA, USA, 2017	Кондиционеры, 30 помещений	Q-learning, neural networks	20-70%
Department of Electrical Engineering, Uppsala University, Sweden, 2019	Горячая вода	Q-learning, regression	15%
Carnegie Mellon University, Pittsburgh, PA, USA, 2018	Тепло, ГВС, кондиционеры	A3C	10-23%

# Варианты схем отопления

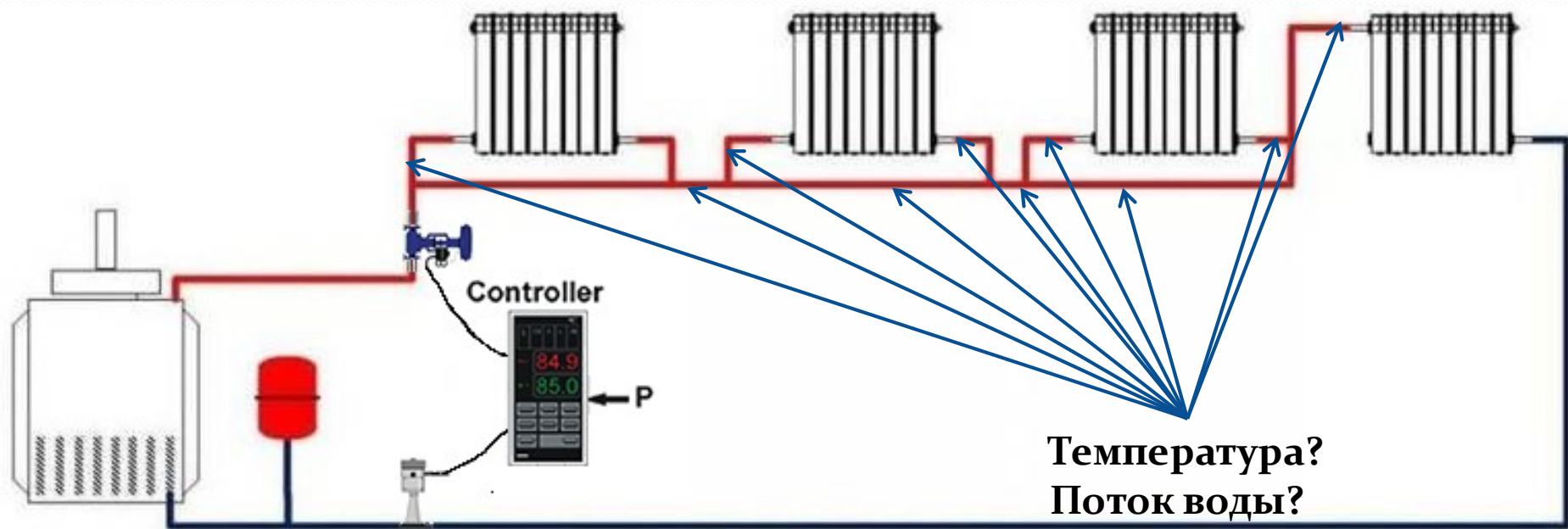


*однотрубная система отопления "Ленинградка"*

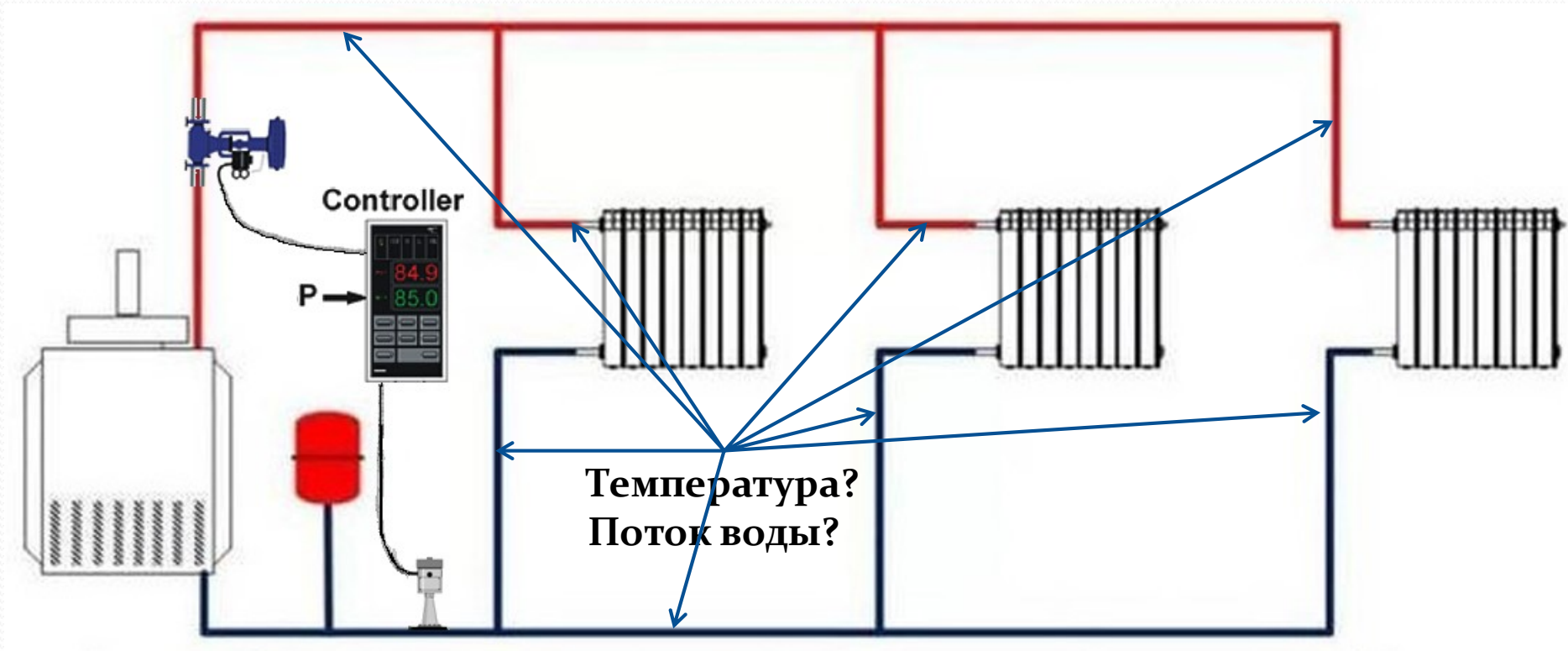


*однотрубная система отопления с принудительной циркуляцией (насосная)*

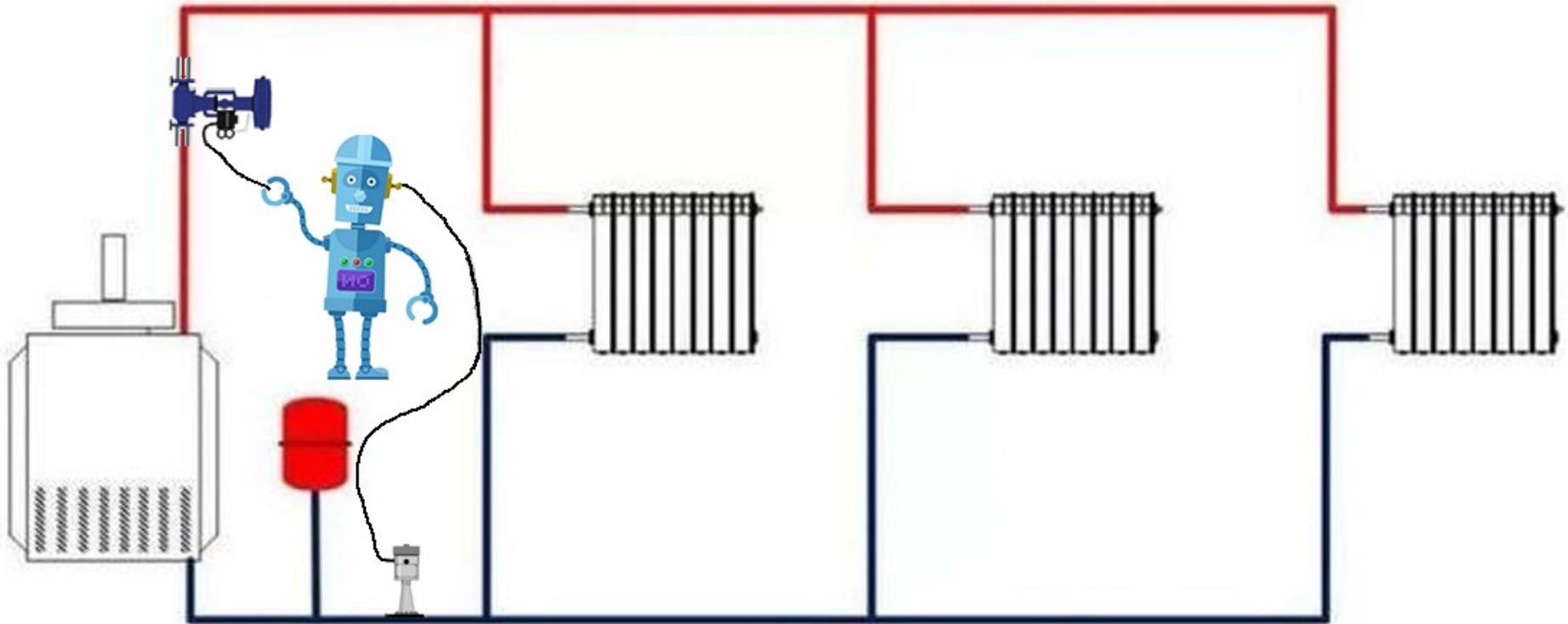
# Современный «неумный» дом



# Современный «неумный» дом



# «Умный» дом v1: RL agent

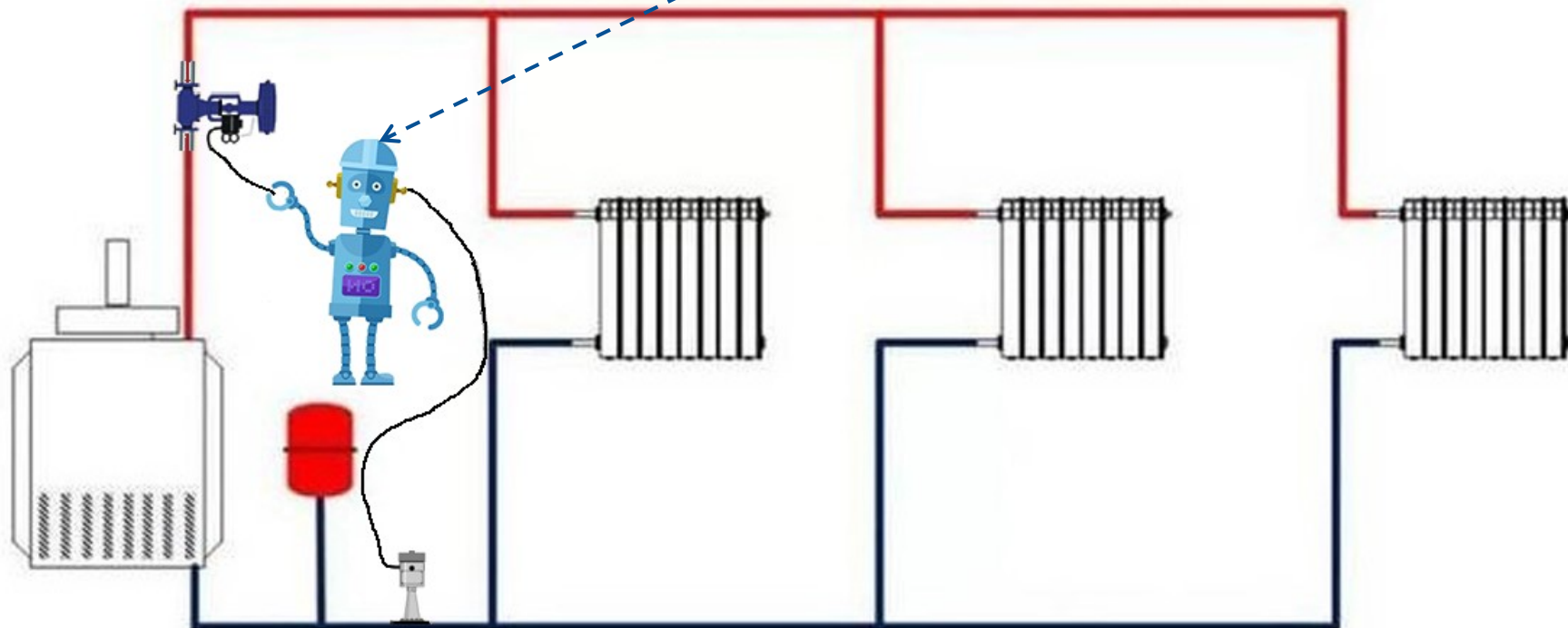


**Экономия  $\approx 10\%$**

# «Умный» дом v2: RL agent + облако

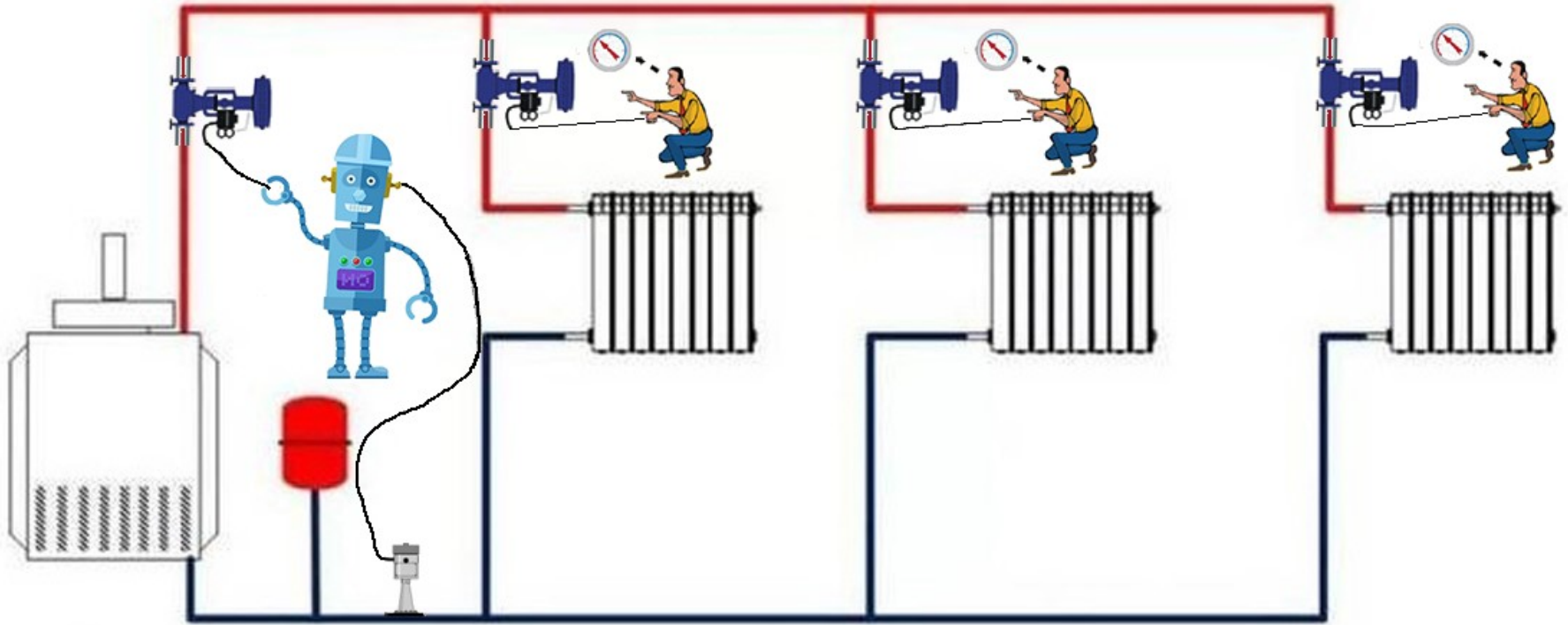
**CLOUD**

Прогноз погоды: температура, влажность, ветер



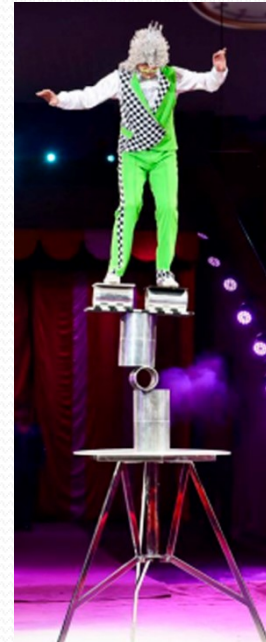
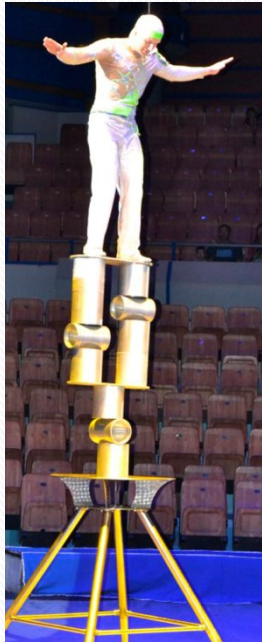
**Экономия ≈ 10-15%**

# «Умный» дом ? : RL агент + «операторы»



**НЕВОЗМОЖНО**

# Образ умного дома



**RL агенты  
+ МНОГО ДАННЫХ**



# Образ современного **неумного** дома

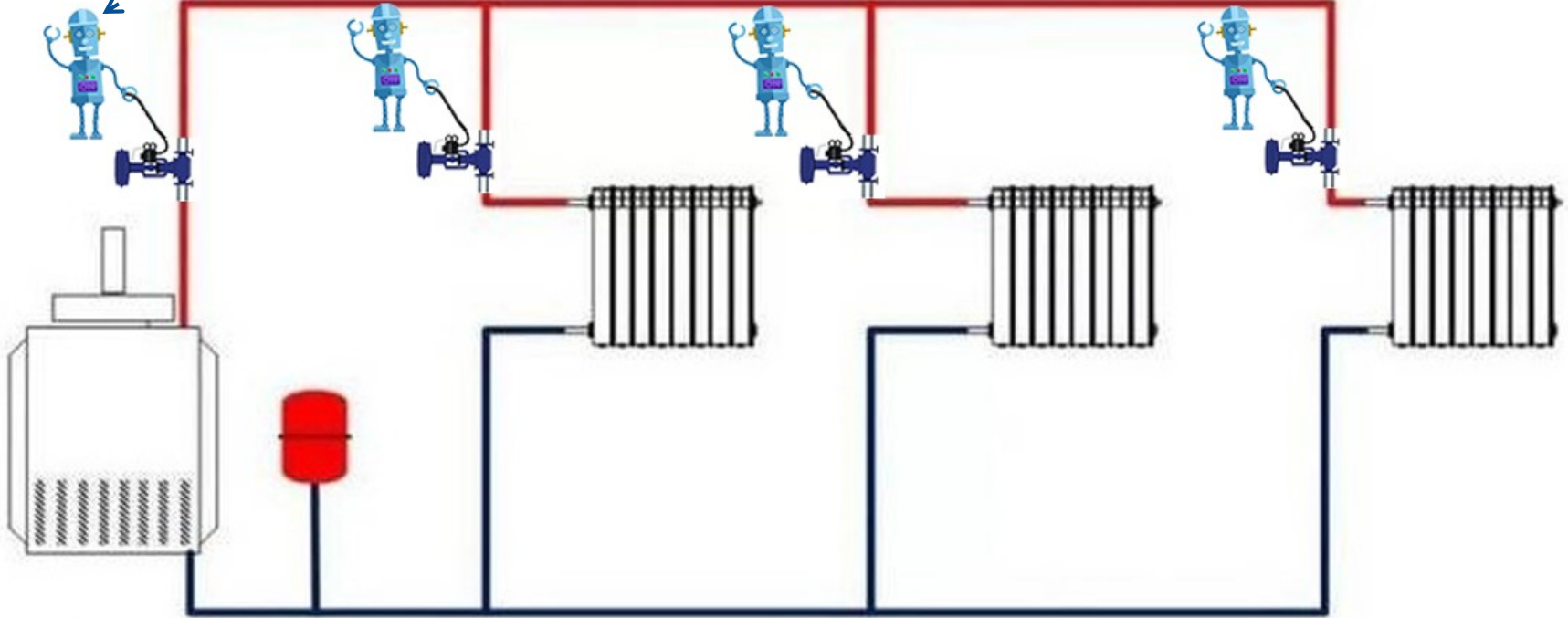


**ПИД регулятор  
+ отсутствие данных**

# Умный дом V3: RL агенты + облако

## CLOUD

Прогноз погоды: температура, влажность, ветер  
Данные от других агентов дома, прочие данные



**Экономия  $\approx$  25-30%**

Неумный

Умный



# Глобальные данные

**На тепло уходит большая часть энергии, производимой на Земле**

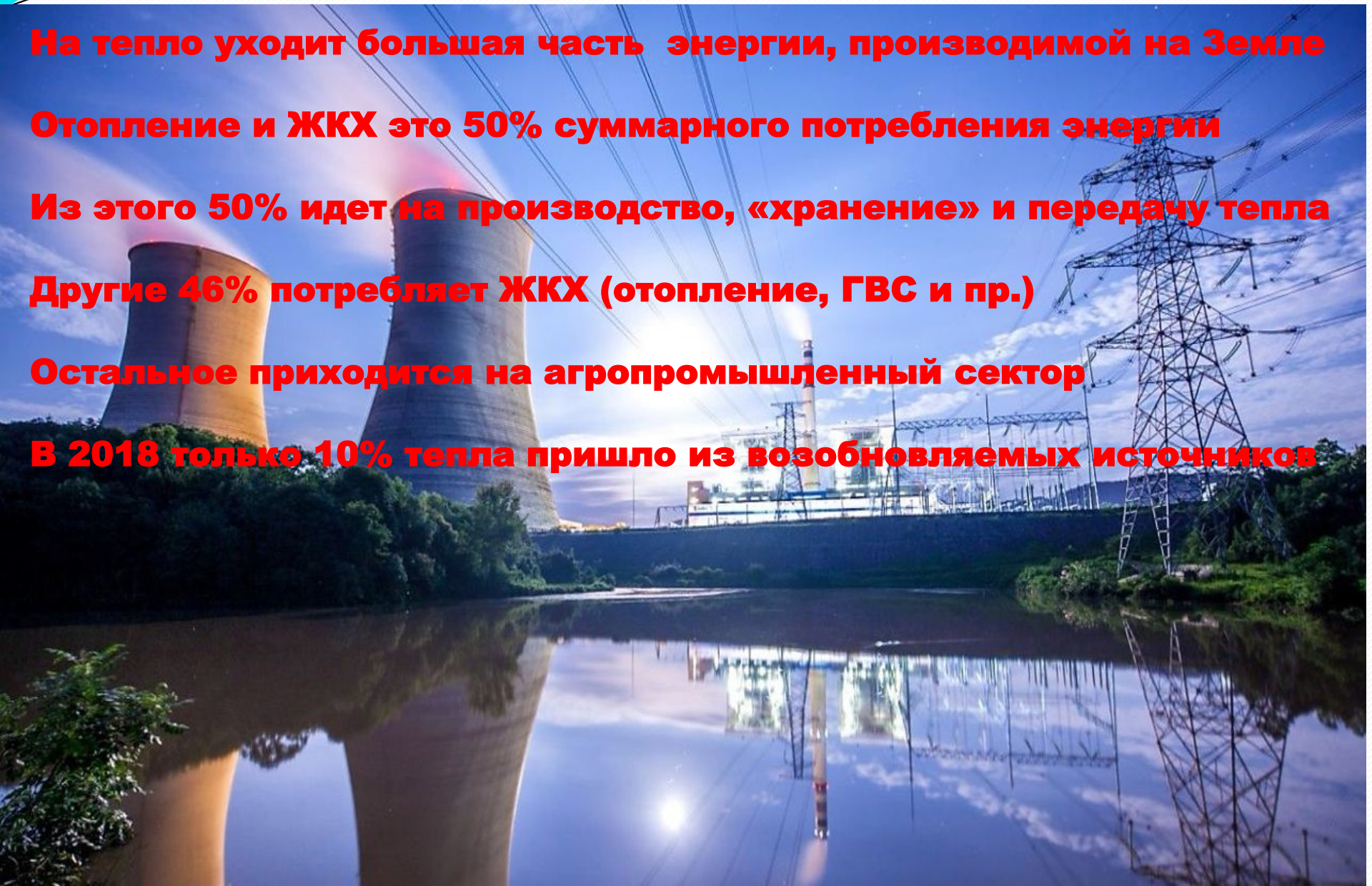
**Отопление и ЖКХ это 50% суммарного потребления энергии**

**Из этого 50% идет на производство, «хранение» и передачу тепла**

**Другие 46% потребляет ЖКХ (отопление, ГВС и пр.)**

**Остальное приходится на агропромышленный сектор**

**В 2018 только 10% тепла пришло из возобновляемых источников**



Контакты:

Геннадий Суворов

руководитель ИИ проектов

SMART4, LLC

<https://www.smart-4.ru/>

тел: +7 925 800 7797

