Kaspersky Industrial Cybersecurity Conference

Kaspersky Machine Learning for Anomaly Detection

Самообучающийся ИИ

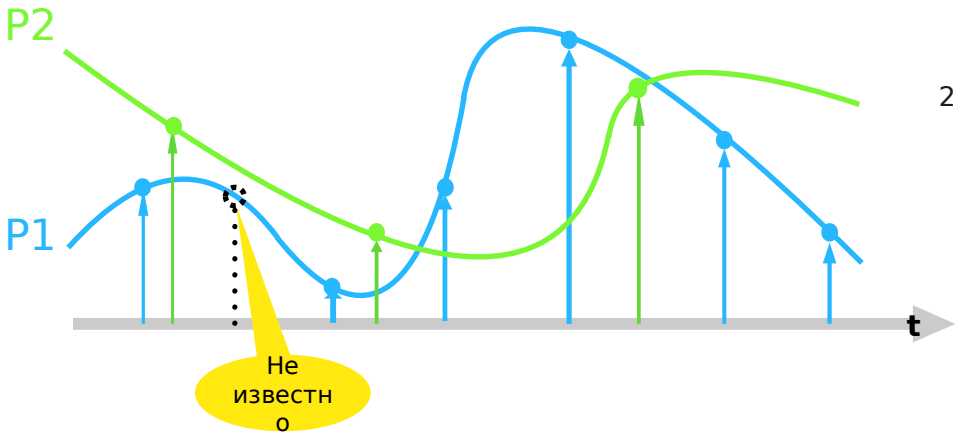
для выявления закономерностей и аномалий в работе предприятия

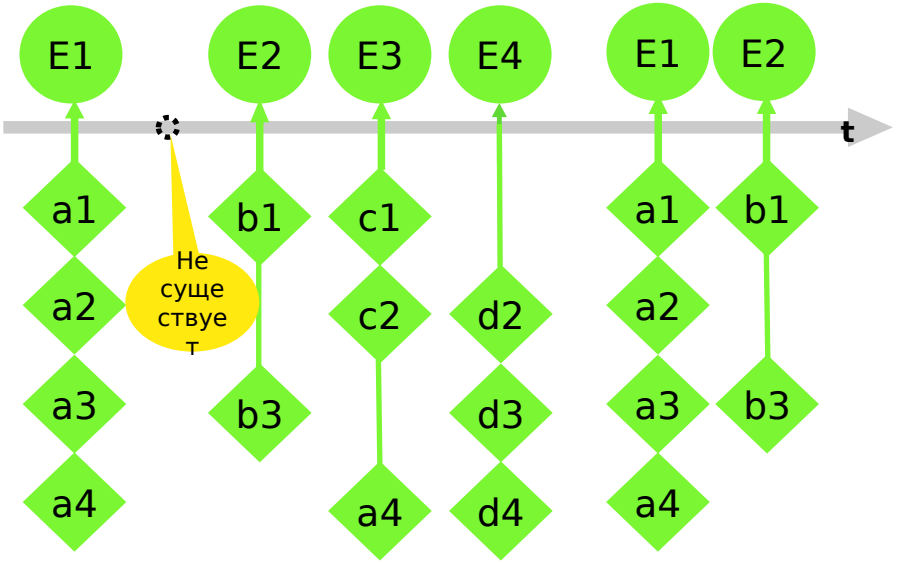
Лаврентьев Андрей

kaspersky

Типы данных

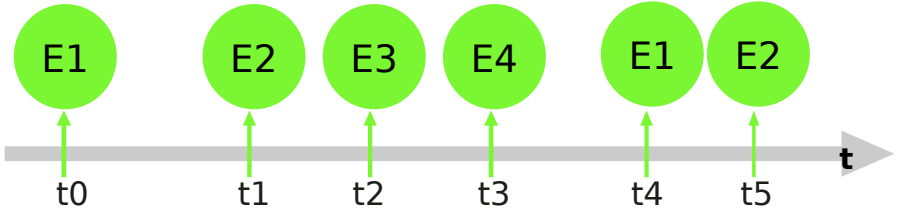
- Непрерывные, числовые — **телеметрия**. Сильно связаны физикой/логикой. Хорошо работает аппроксимация.
- Дискретные, категориальные — **события** (время и атрибуты). Атомарные — нет значения между двумя событиями. Важны порядок и ин
- Далее говорим про события





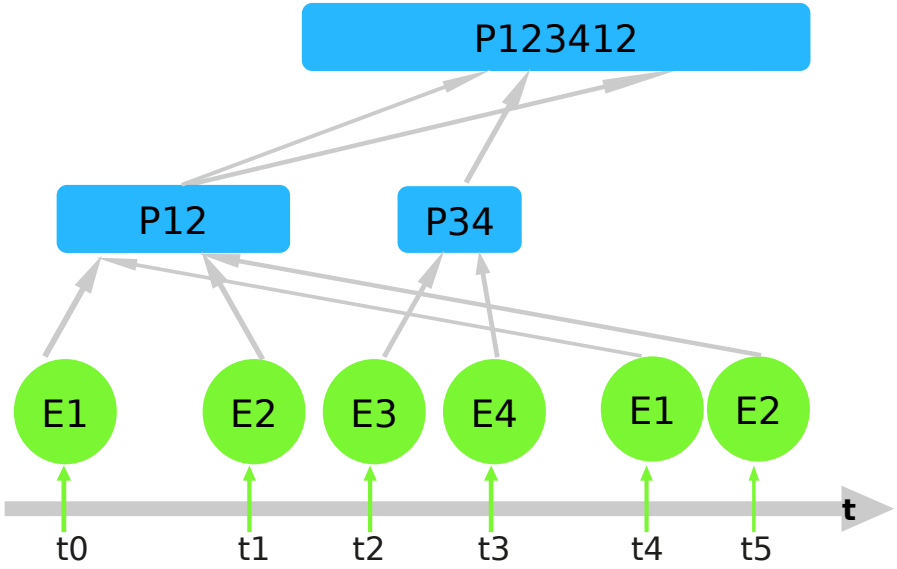
Норма? Аномалия?

- Норма - штатный порядок и интервалы
- Аномалия - нарушение нормы
- А если объект сложный, и процесс не периодический? Как заставить систему «сказать» , что есть норма? и в чем у аномалии отклонение от нормы?



Норма? Аномалия?

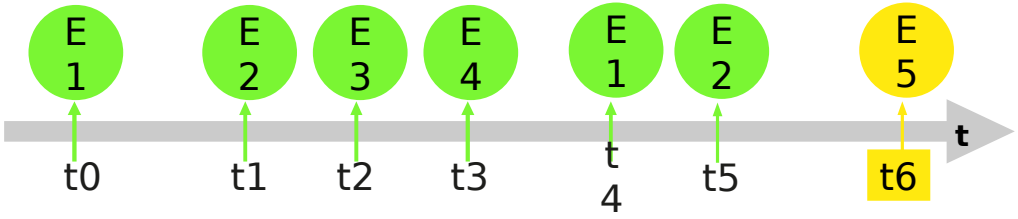
- Норма - штатный порядок и интервалы
- Аномалия - нарушение нормы
- А если объект сложный, и процесс не периодический? Как заставить систему «сказать» , что есть норма? и в чем у аномалии отклонение от нормы?



- Норма - не только последовательность, это **структура** последовательности — **иерархия** вложенных паттернов — из коротких паттернов строятся более длинные и т.д. Так же как из букв строятся морфемы, из морфем слова и т.д.

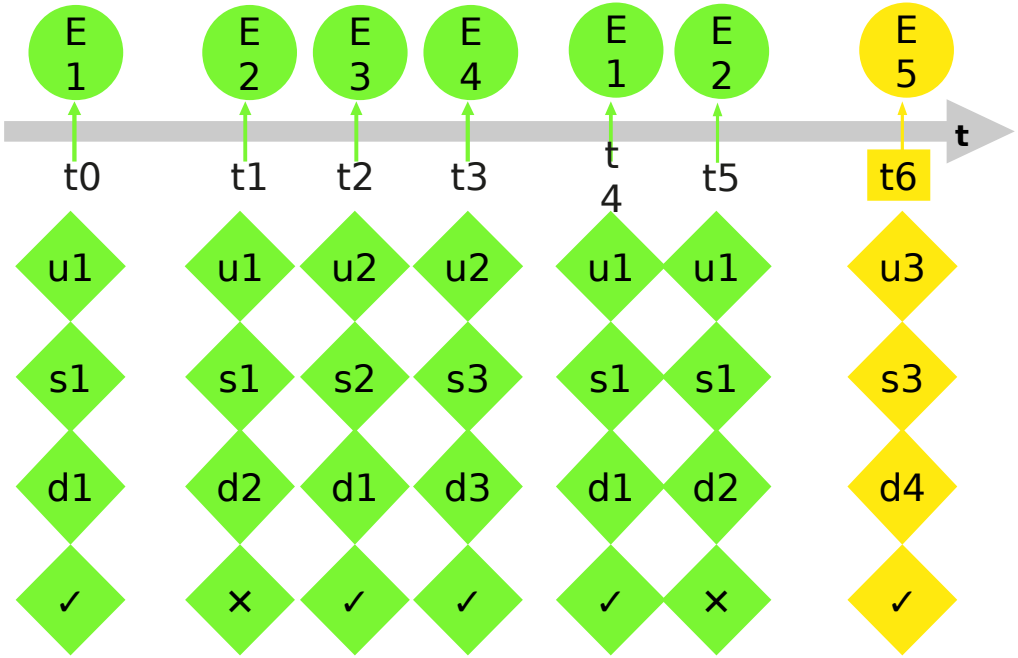
Нейросеть - но какая?

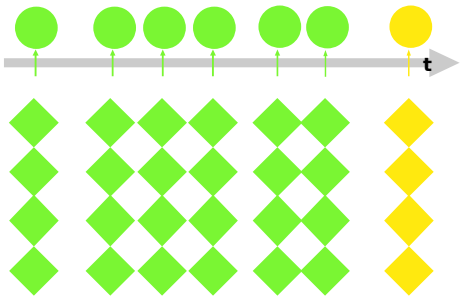
- Должна **постоянно учиться** — т.к. нет исчерпывающих примеров со всем разнообразием паттернов



Нейросеть - но какая?

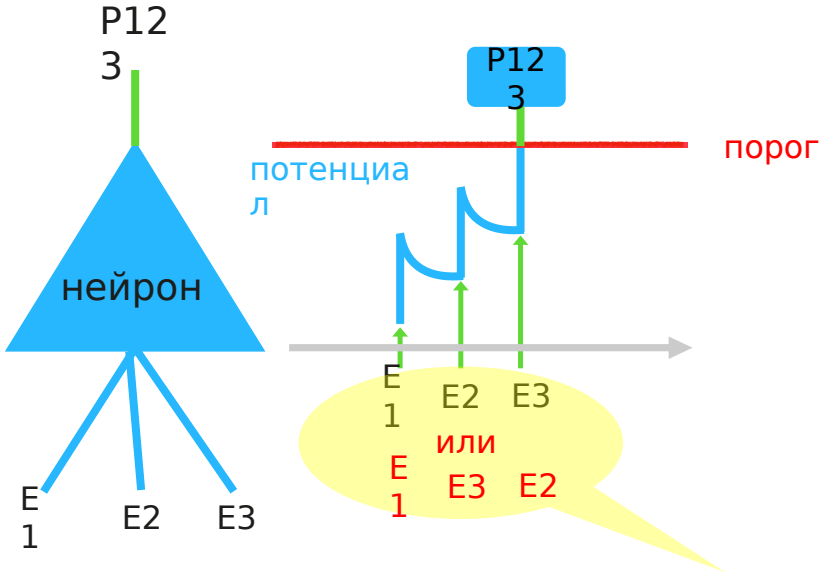
- Должна **постоянно учиться** — т.к. нет исчерпывающих примеров со всем разнообразием паттернов
- Должна обрабатывать событийные данные = **время** + **значения набора параметров**





Нейросеть - но какая?

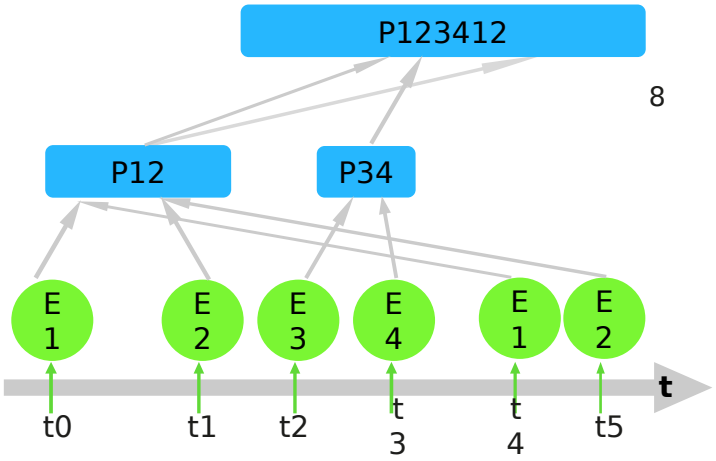
- Должна **постоянно учиться** — т.к. нет исчерпывающих примеров со всем разнообразием паттернов
- Должна обрабатывать событийные данные = **время** + **значения набора параметров**
- **? Решение** — импульсные сети, нейроны которых активируются импульсами (событиями)
- Необходимо видеть что "видит" сеть !!!

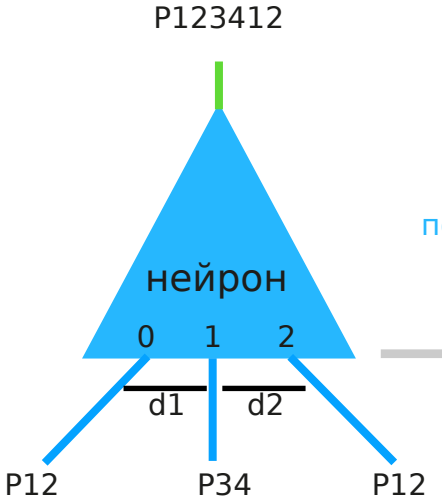


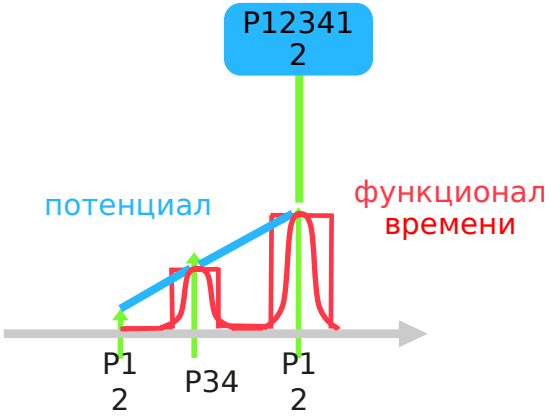
- Необходимо что бы нейрон различал порядок и временные интервалы между импульсами!

Нейросемантическая сеть

- Предложена в 1990 В. Бодякиным, развита автором для учета интервалов времени
- **Послойная** — на каждом слое пытается "подобрать" паттерны (нейроны) по данным, представленных паттернами (нейронами) предыдущего слоя — так что бы **общий словарь** паттернов на всех слоях в битах был бы **наиболее компактным**







"Чувствует" семантику

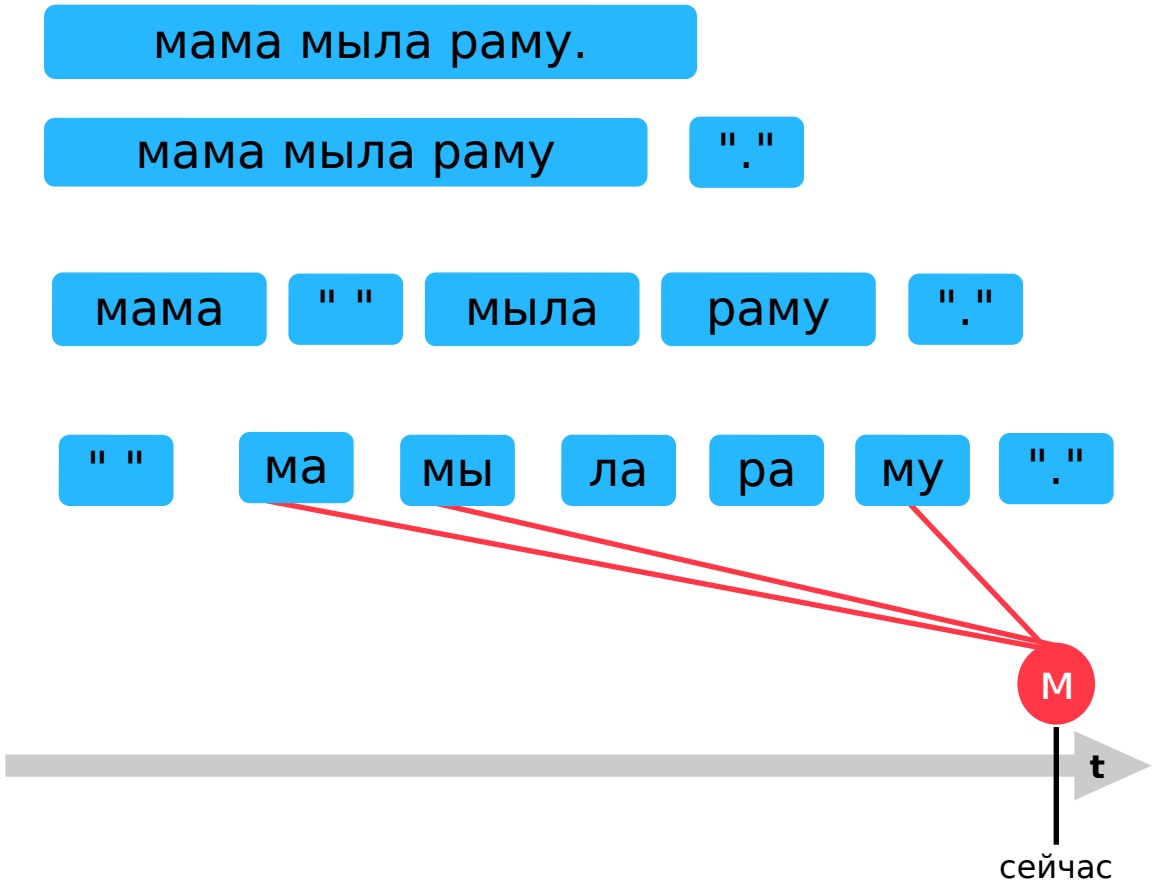
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

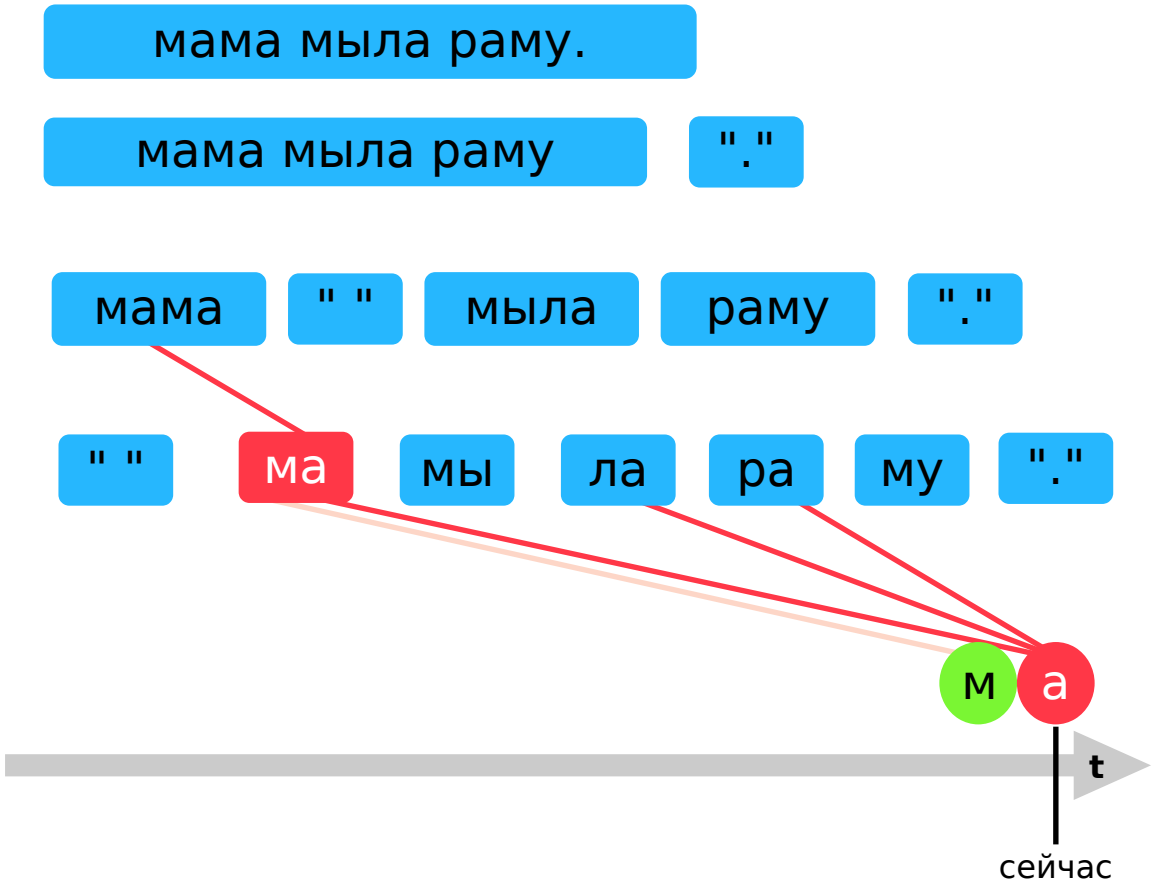
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

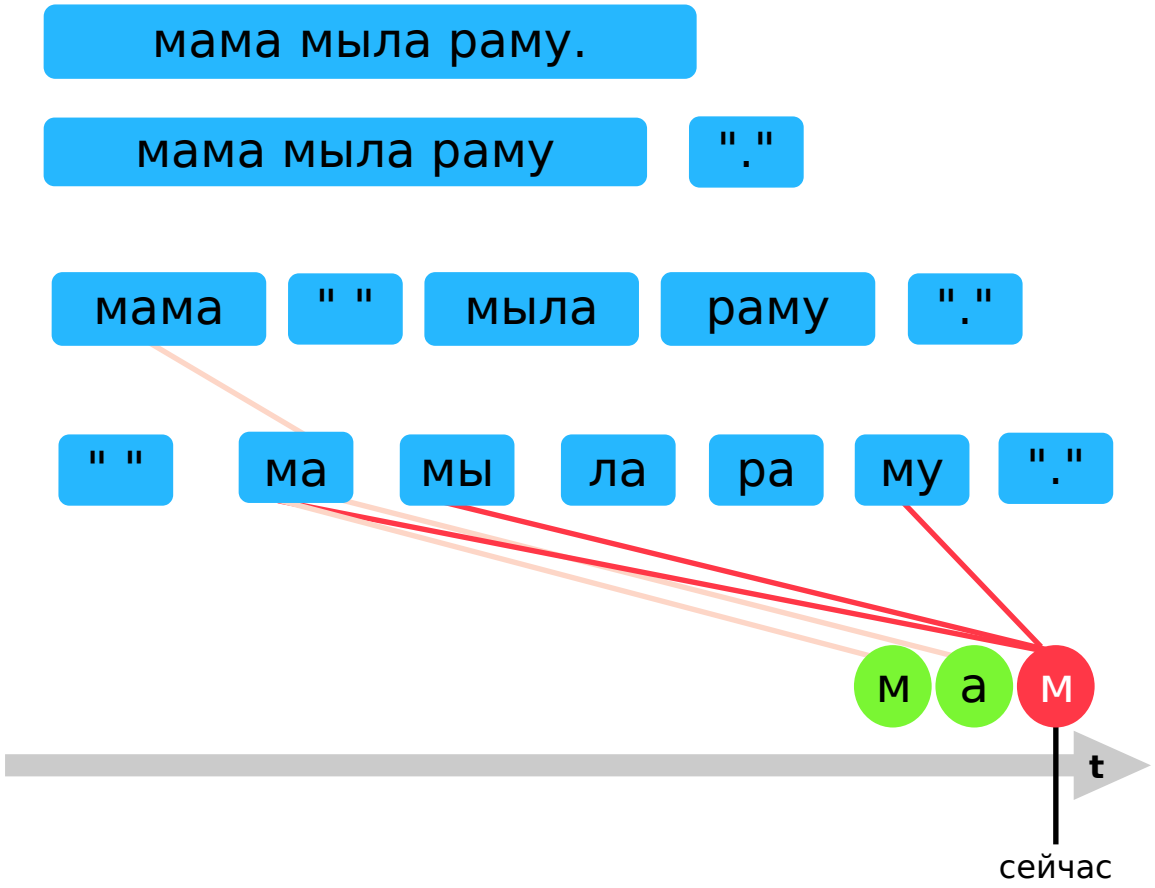
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

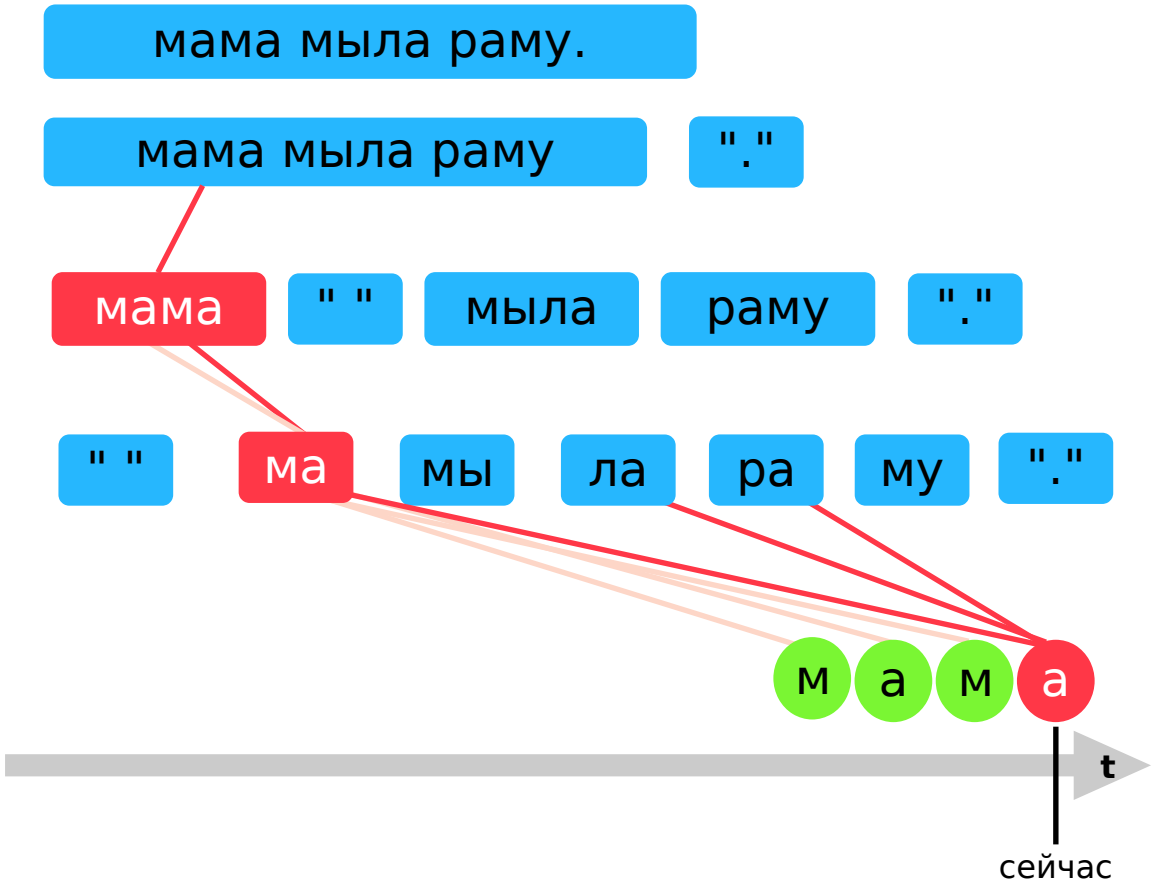
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

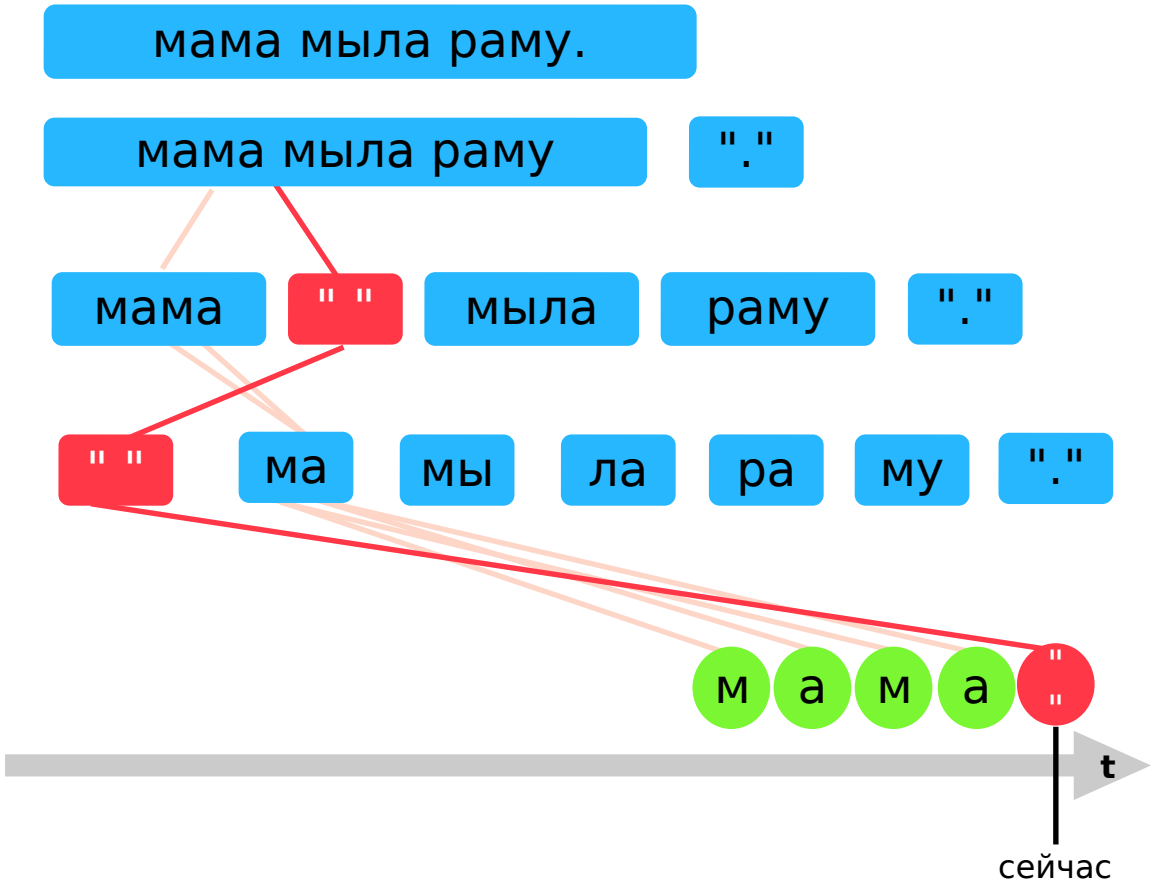
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

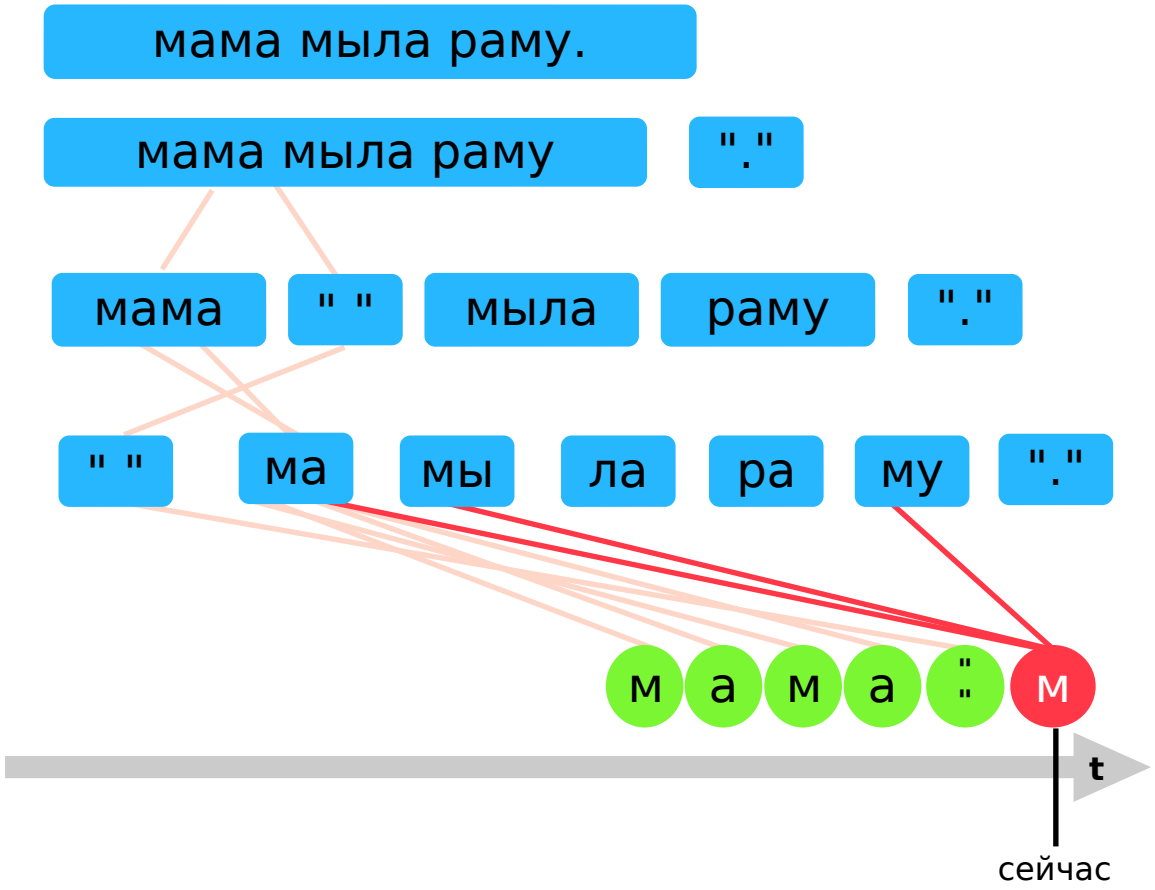
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

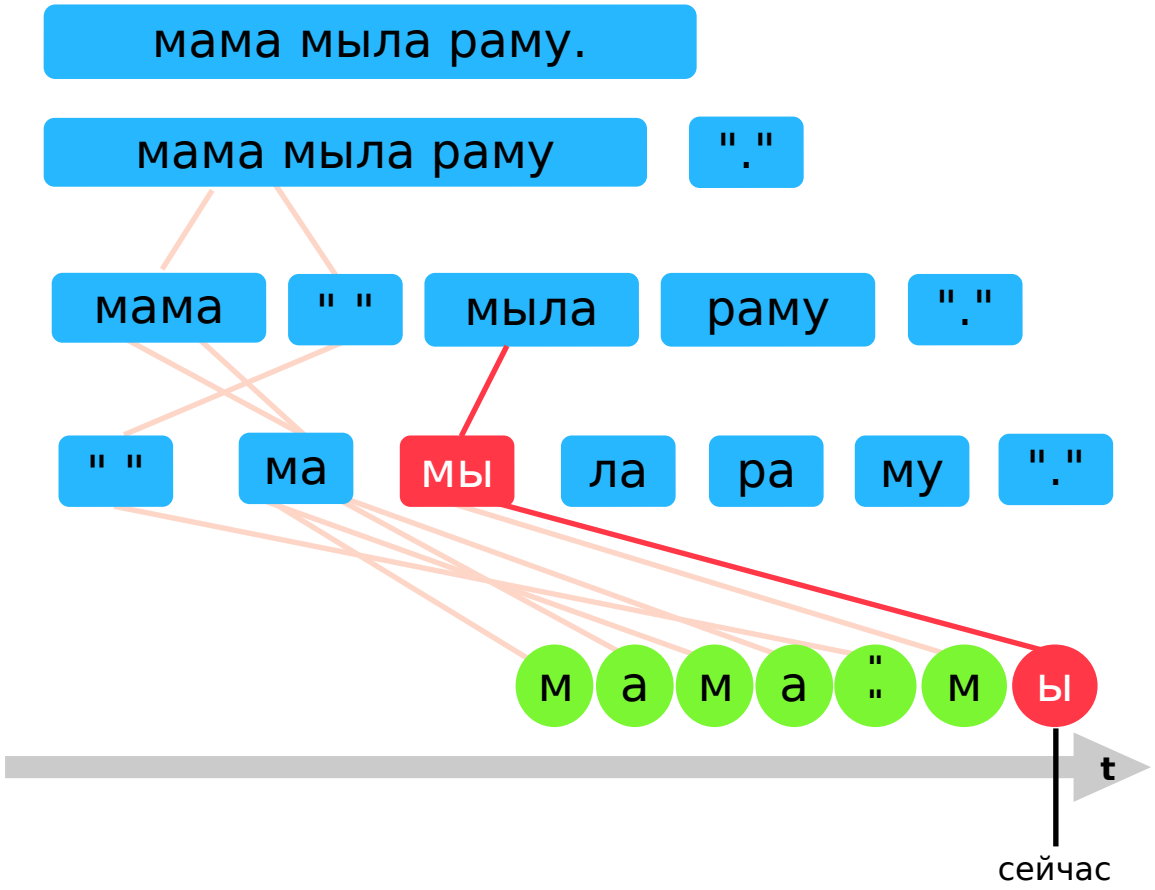
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

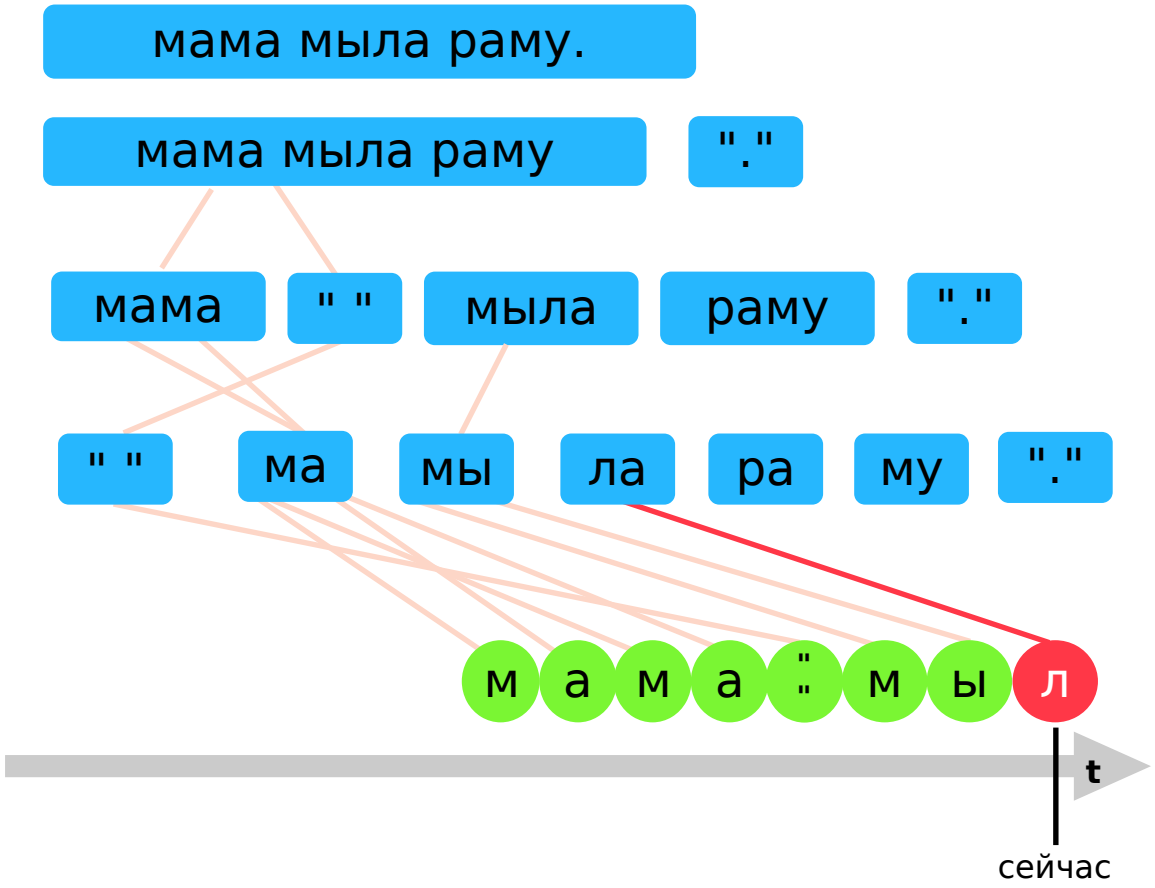
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

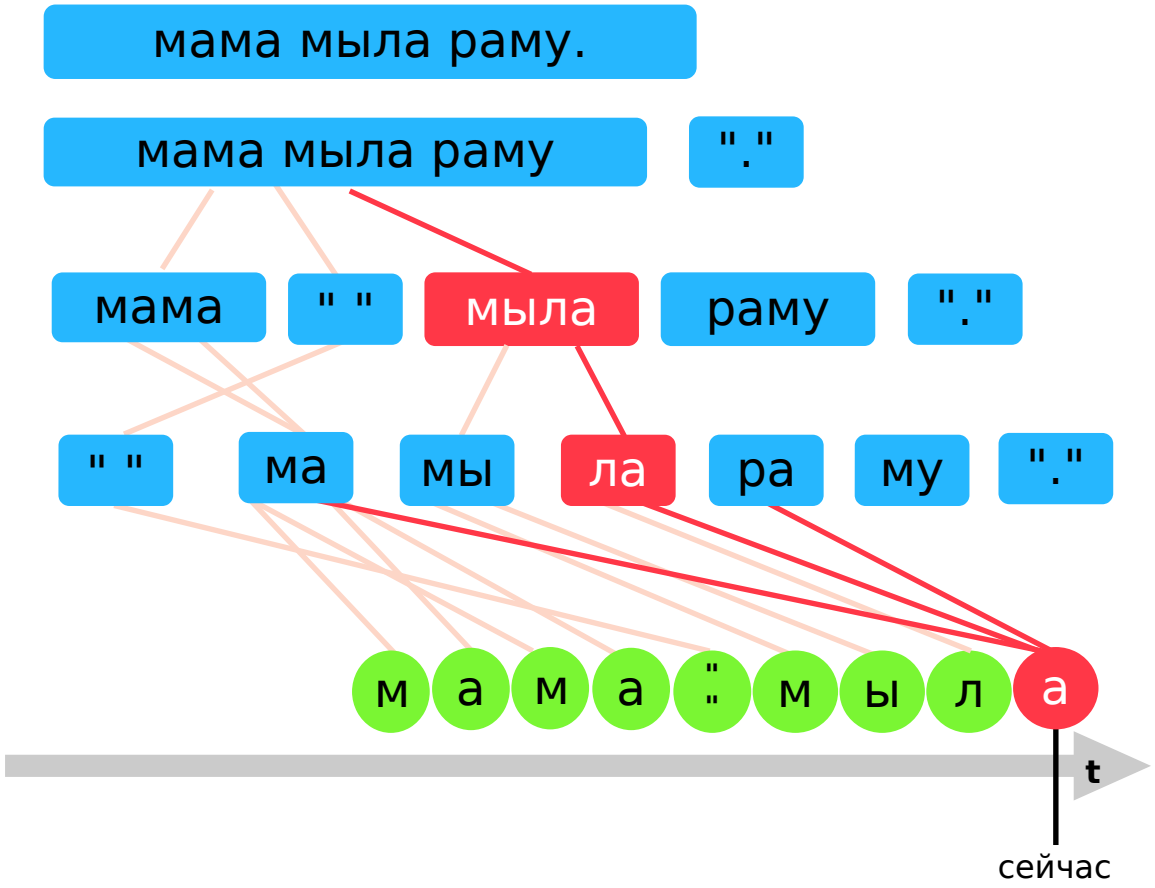
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

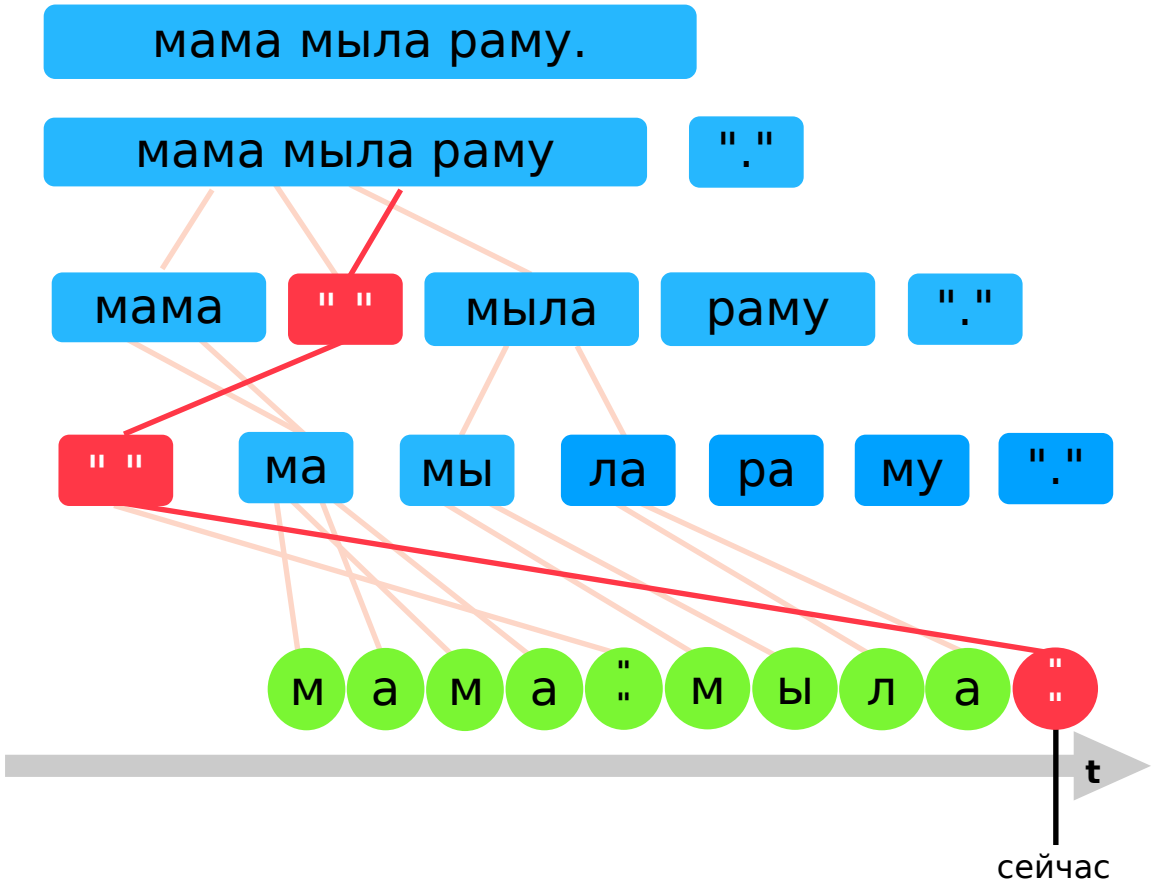
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

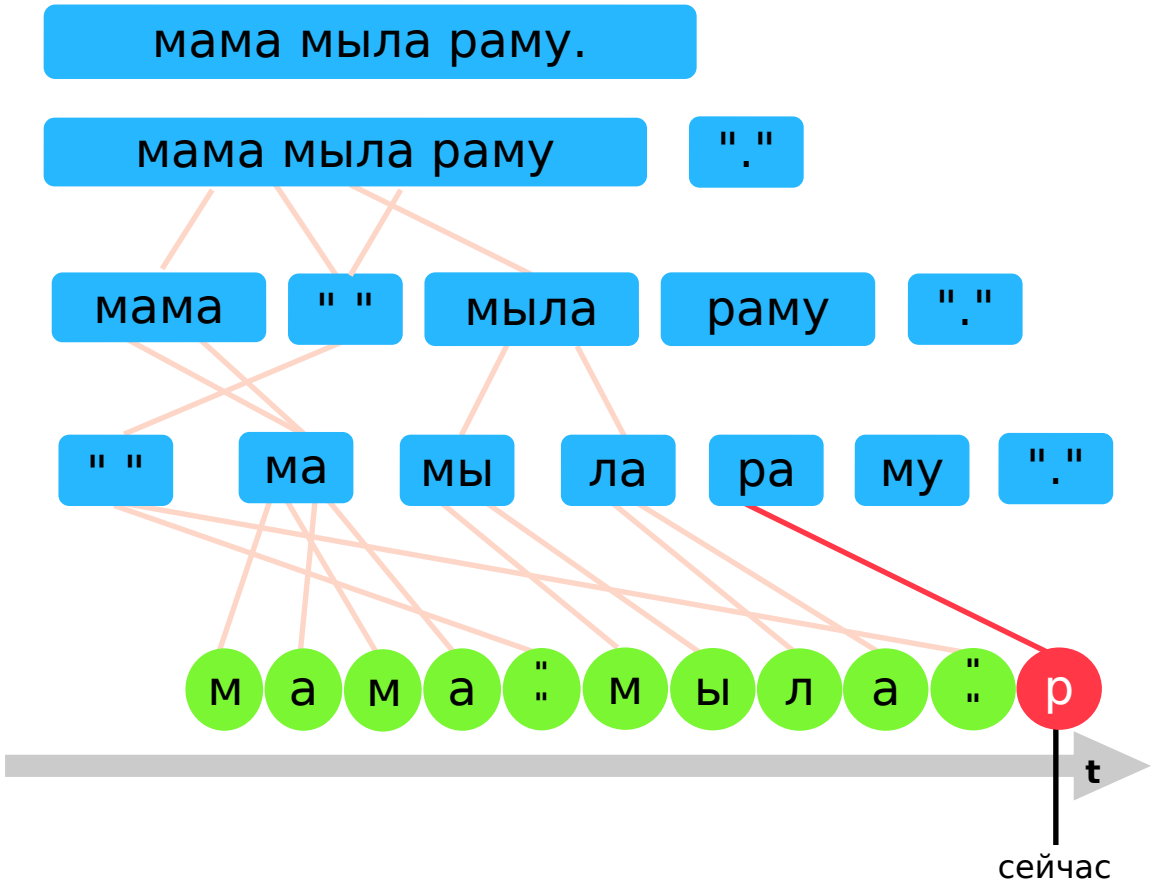
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

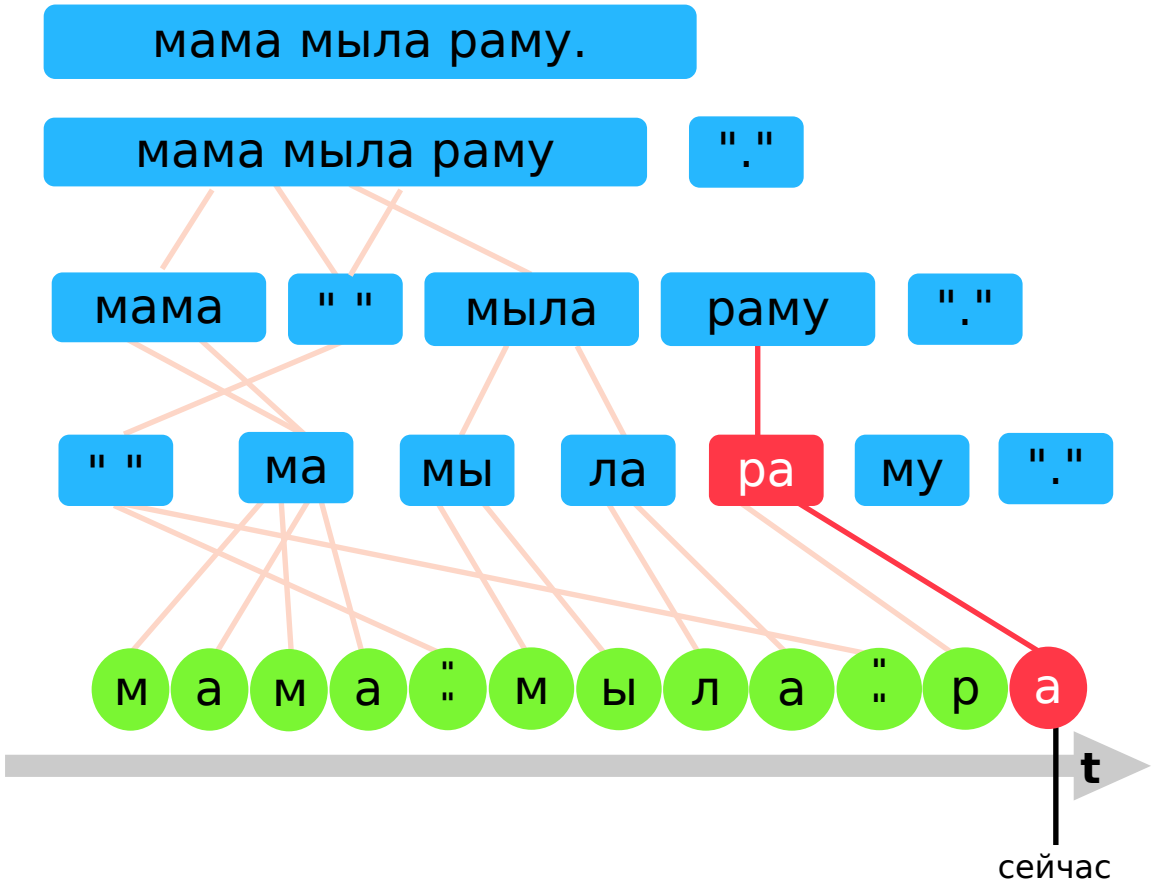
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

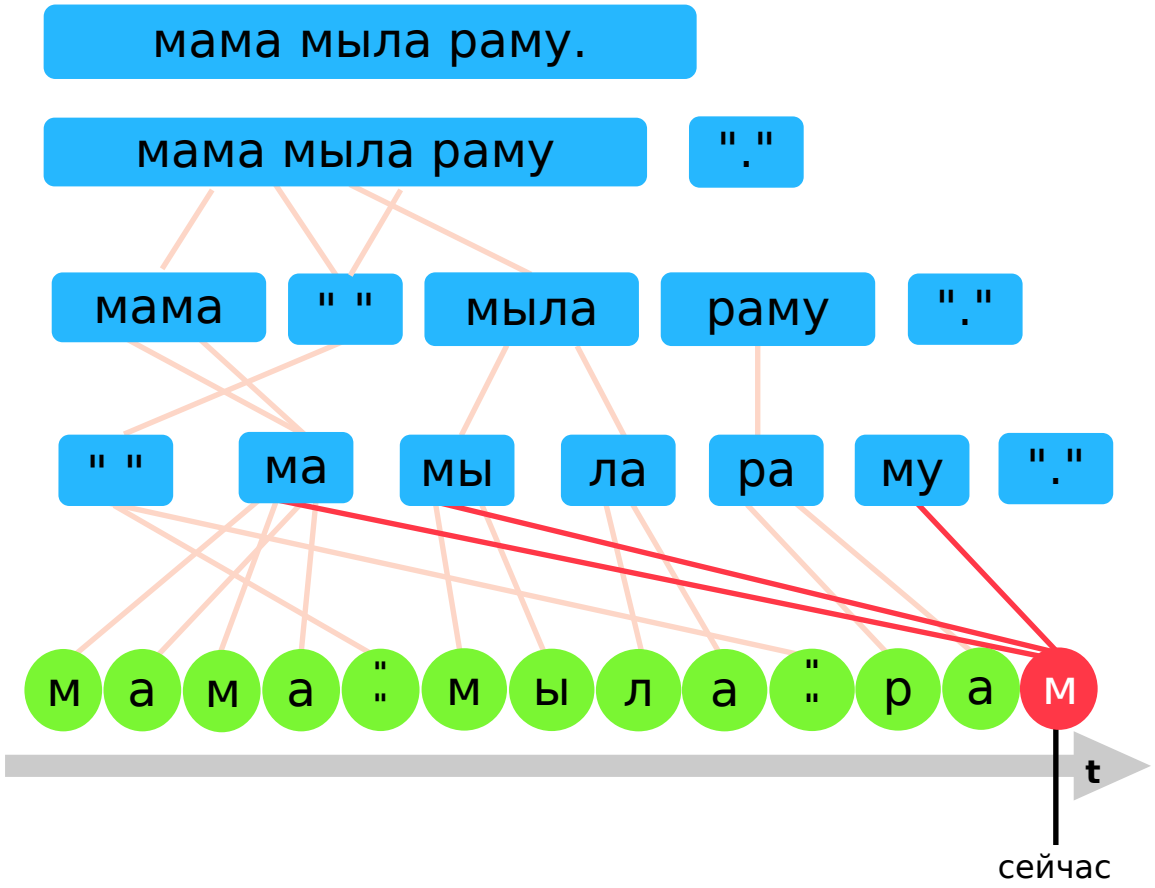
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

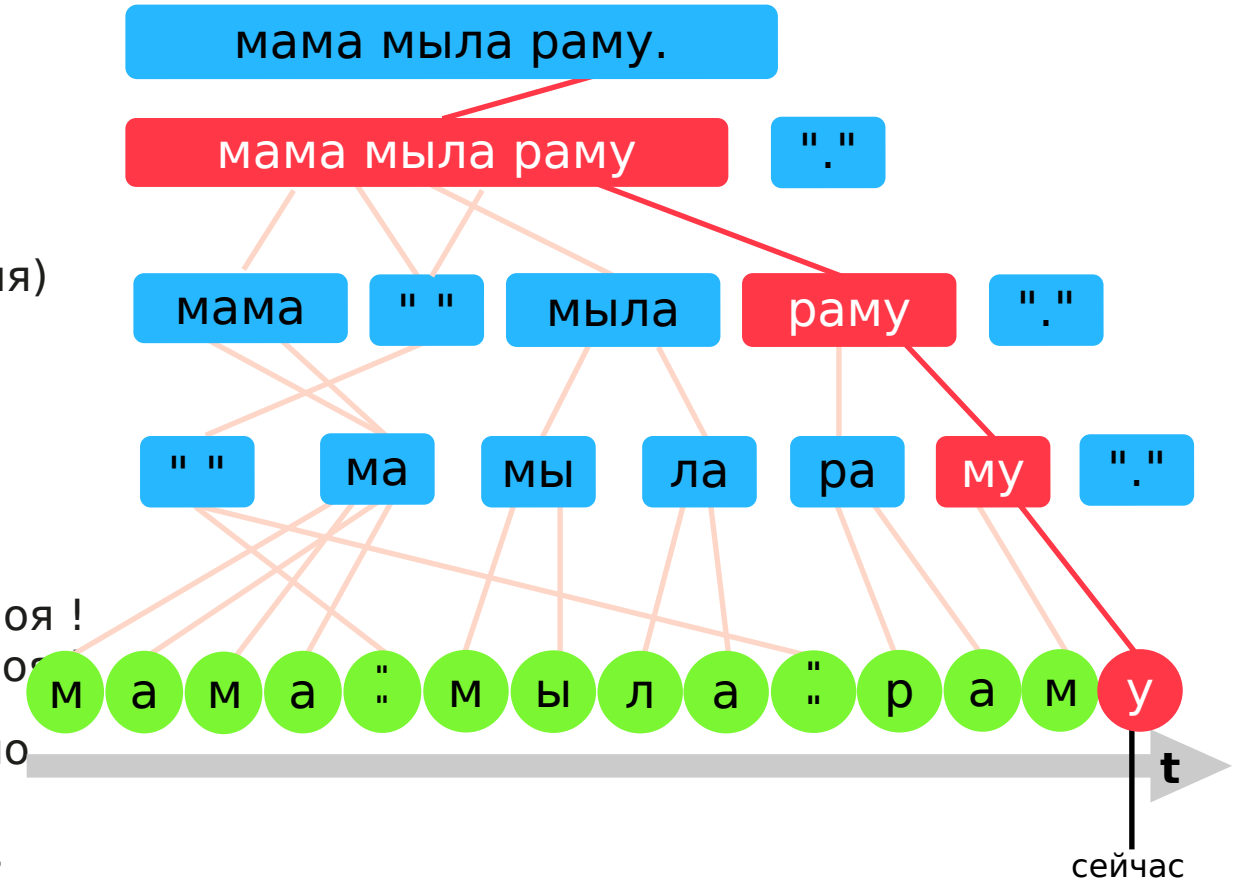
  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

- Слои паттернов:
  - 0 — буквы (vs события)
  - 1 — морфемы
  - 2 — слова
  - 3 — фразы
  - 4 — предложения
- **Но:**
  - " " — паттерн 2-го слоя !
  - "." — паттерн 3-го слоя
- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



"Чувствует" семантику

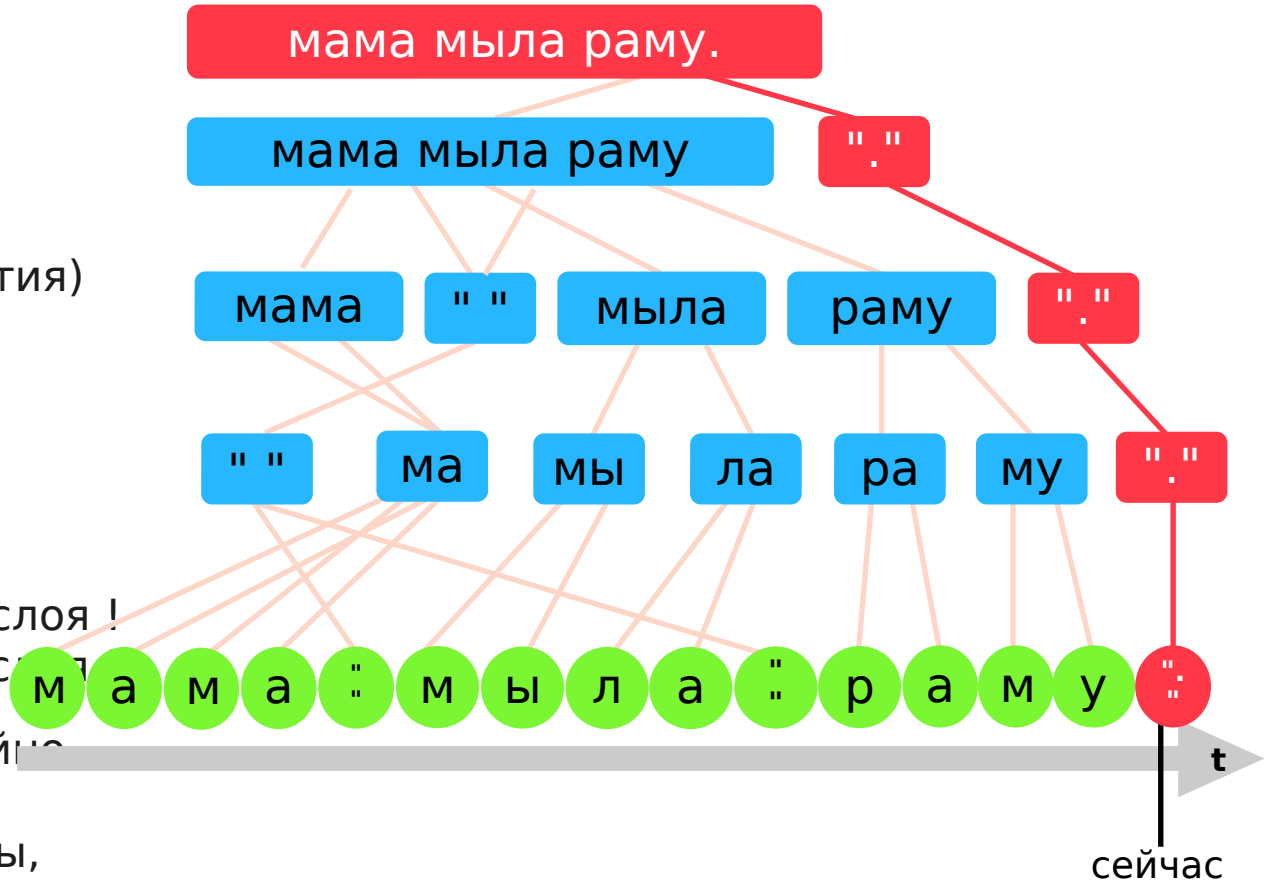
- Слои паттернов:

  0 — буквы (vs события)
  1 — морфемы
  2 — слова
  3 — фразы
  4 — предложения

- **Но:**

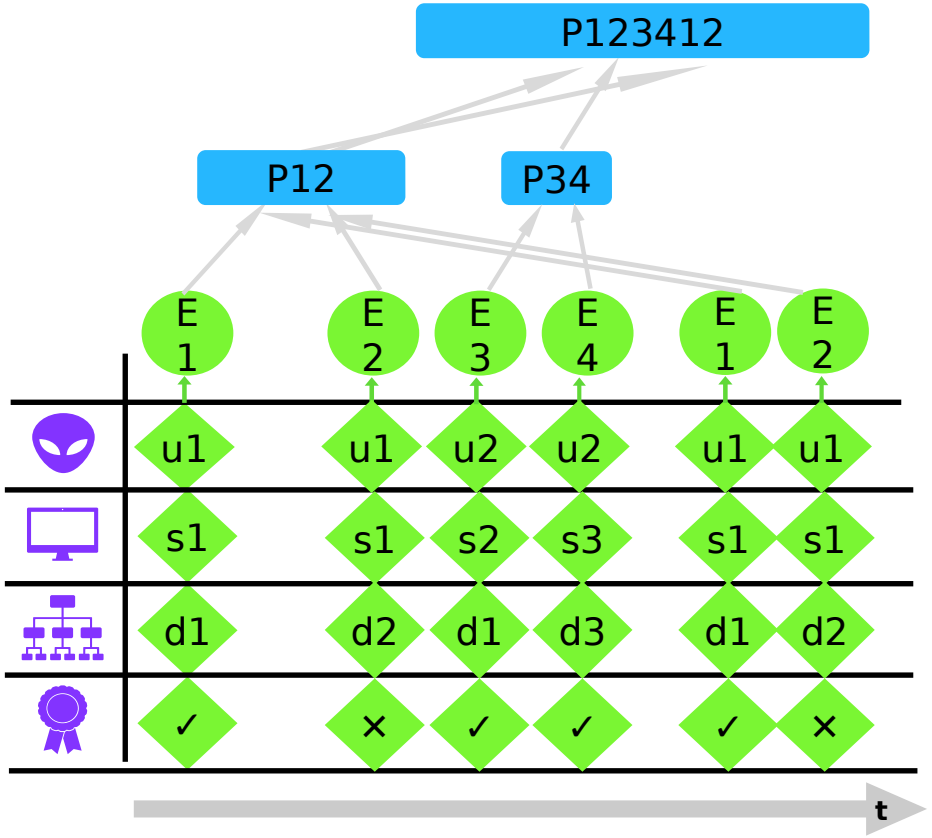
  " " — паттерн 2-го слоя !
  "." — паттерн 3-го слоя !

- Сеть на потоке послойно "ищет" наиболее употребимые морфемы, слова, фразы



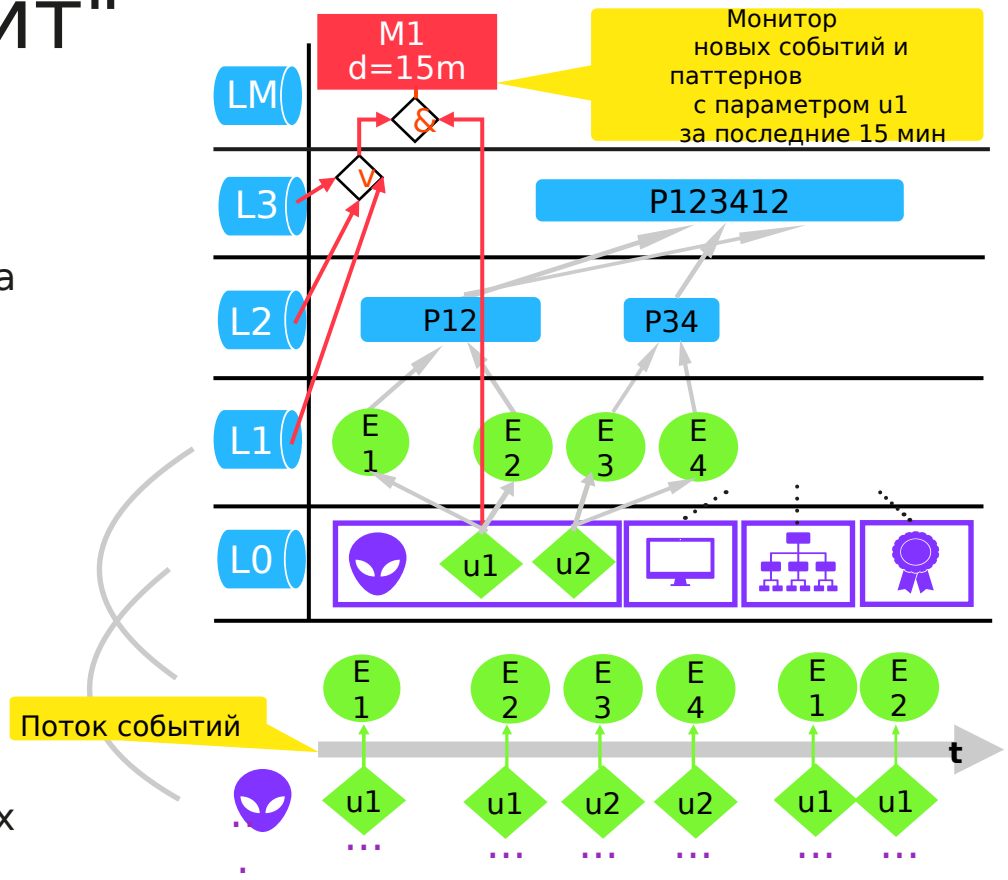
Что "видит" сеть?

- **Входные каналы.** Аналогия — 5 чувств человека, но у сети этих "чувств" может быть гораздо больше.
- Данные канала – **категориальные** значения параметров с **открытым словарем**
- **Событие** = набор значений параметров, пришедших одновременно по разным каналам
- **Направления внимания** — выделяют из потока определенные события, и уже среди них сеть выявляет паттерны, что уменьшает информационный шум от случайных корреляций несвязанных событий.
- Сеть "видит" **создание и активацию нейронов**, которые соответствуют на потоке значениям параметров



Увидеть что "видит" сеть

- Сеть "общается" с человеком через **выходные каналы — мониторы**
- Создаются «на лету», как специальные **нейроны-мониторы** на специальном слое сети Lm
- Связанны с сетью по нечетким критериям, заданным человеком в терминах:
  - интервалов времени,
  - значений атрибутов событий,
  - направлений внимания,
  - частоты повторяемости
- **Интерпретация — полностью прозрачна** — основана на разреженных (sparse) упорядоченных связях нейронов



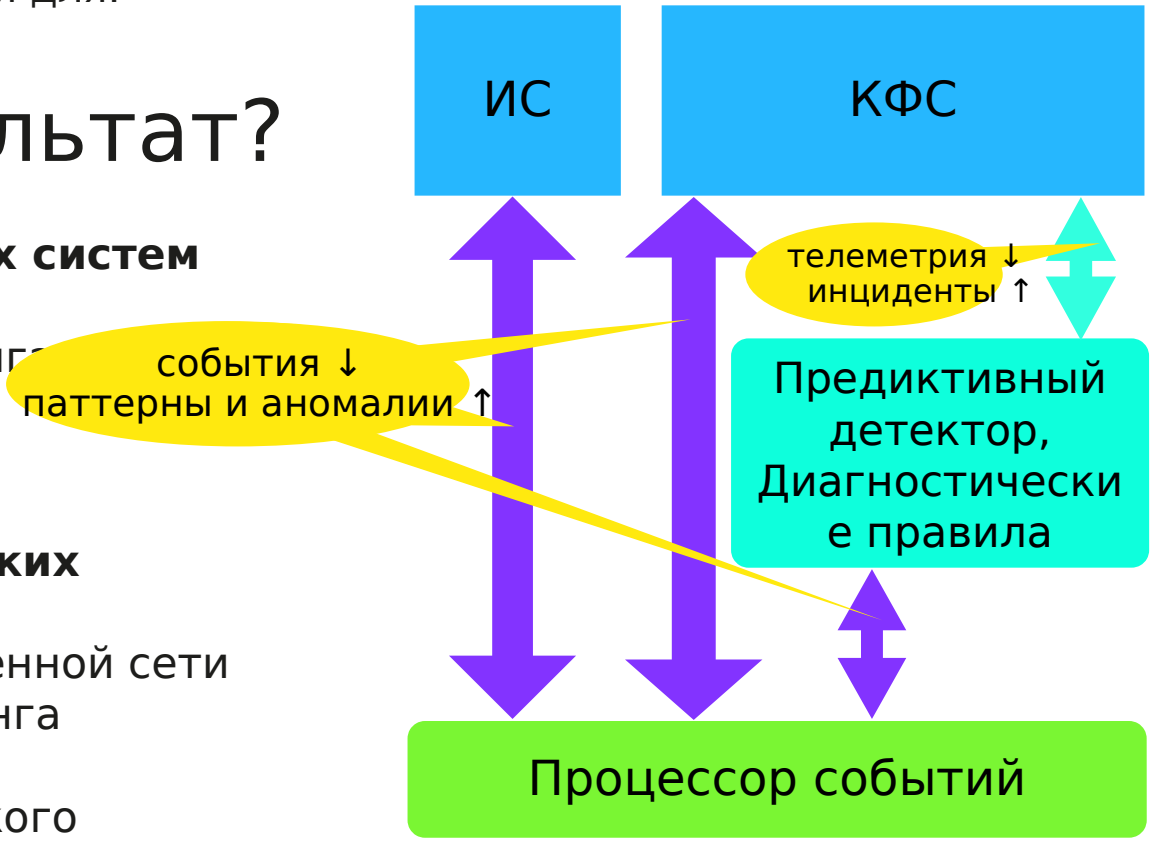
Какой результат?

ИС — информационных систем

- события сети
- события мониторинга пользователей
- SIEM, KUMA, ...

КФС — кибер-физических систем:

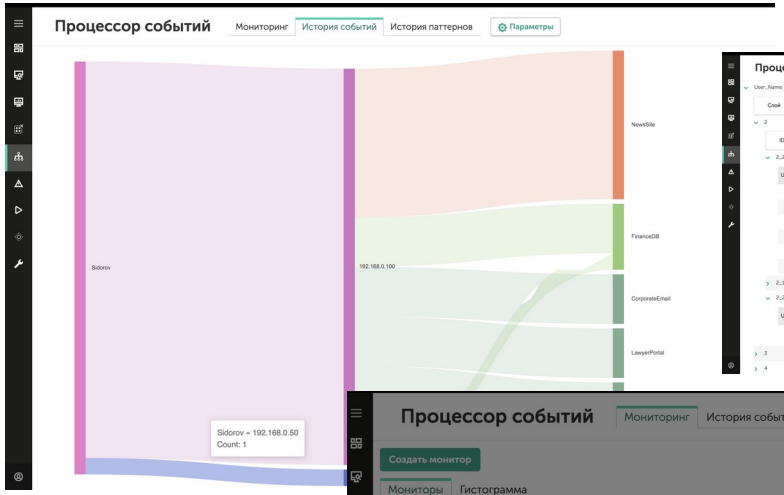
- события промышленной сети
- события мониторинга пользователей
- контроль физического доступа
- аварийные события

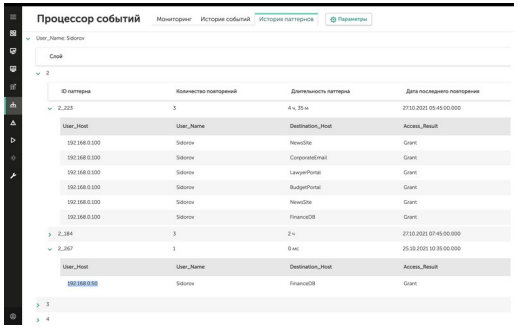


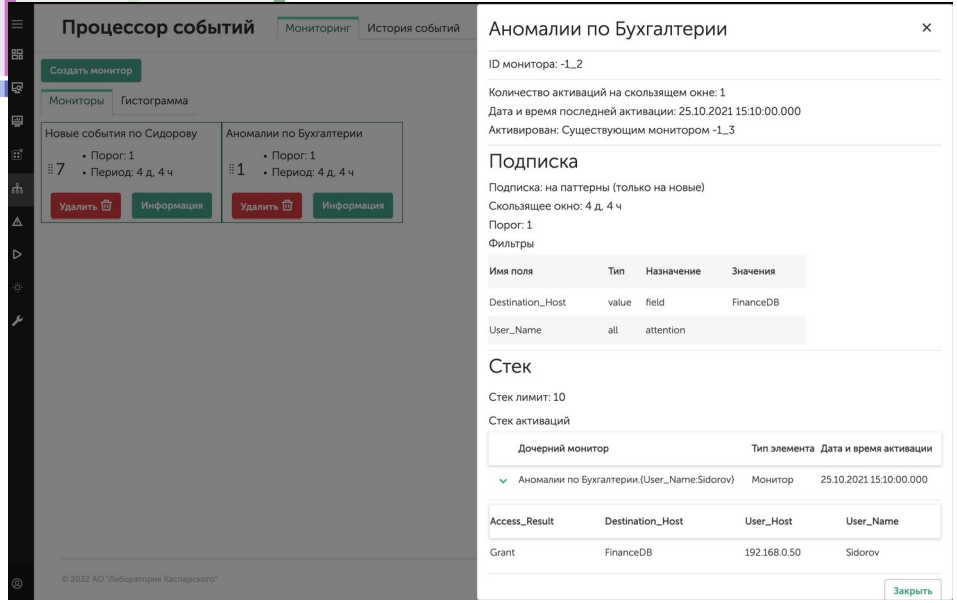
Примеры

- Мониторы
- События
- Паттерны

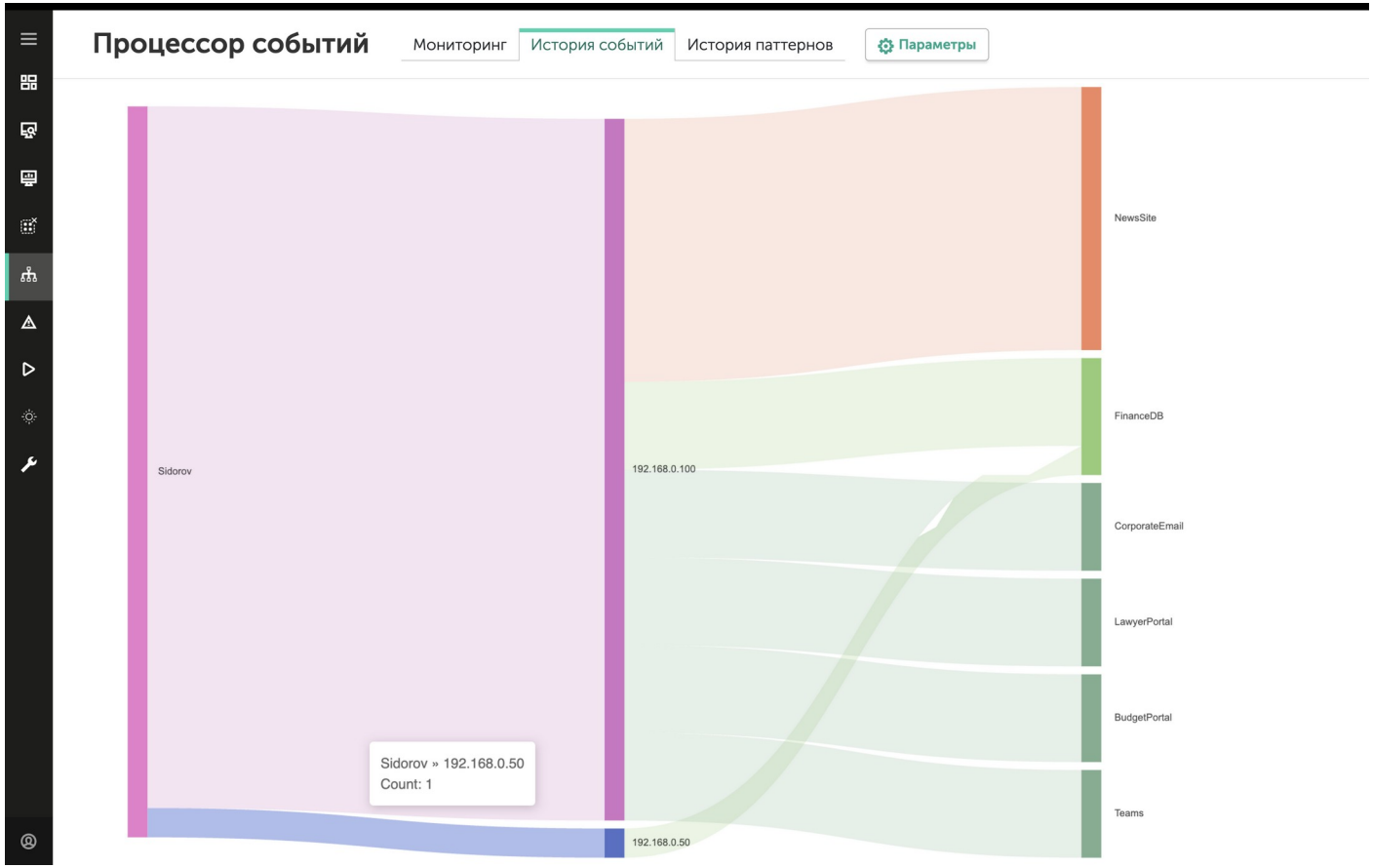
- Аномалии на сервере бухгалтерии







Пример: Граф истории событий в разрезе параметра Пользователь



Самообучающийся ИИ

для выявления закономерностей и аномалий в работе предприятия

уже встроен в

MLAD

Kaspersky®
Machine Learning
for Anomaly Detection

mlad.kaspersky.com
mlad@kaspersky.com