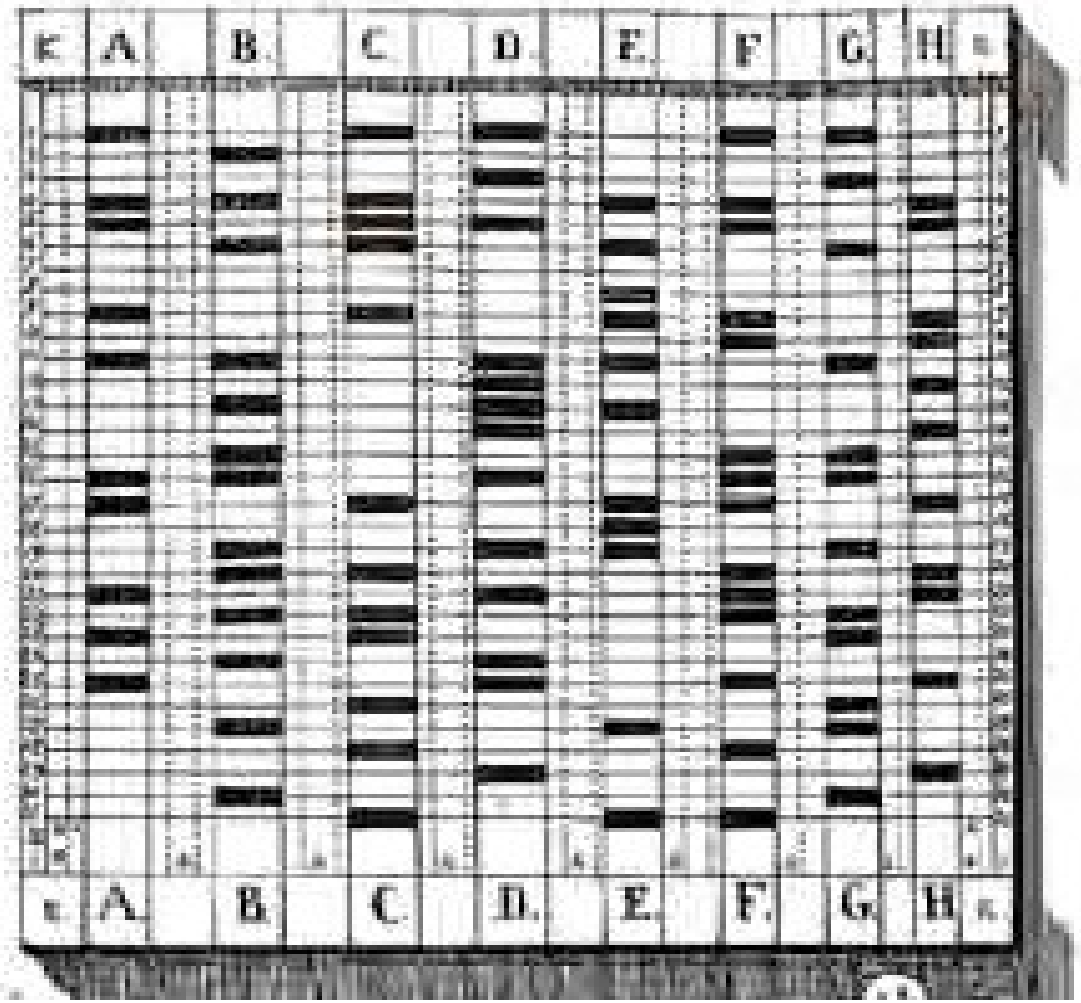
Лекция 1.

Проблемы ИБ решаемые ***экспертными системами***. Основыне понятия. Методы оценки качества экспертных систем.

Павел Владимирович
Слипенчук
PavelMSTU@stego.su
ИУ-8

Москва
09.09.2016

Гомеоскоп Корсáкова (1832)



Семён Николаевич Корсаков (1787-1853)

История. Заметки

“Паскалина”
суммирующая машина
Блеза Паскаля
(1642)

Разностная машина
Иоганна Мюллера
(1788, не построена)

Гомеоскоп
Семена Николаевича
Корсакова
(1832)

Разностная машина
Чарльза Бебиджа
(малая 1822, построена)
(большая 1849,
построена в 2000)

Фашистский
антикварный
проект
(первый гипертекст)
(1937-1943)

Советский физиолог
Пётр Кузьмич Анохин
Формулирует понятие
“*обратной связи*”
(1935)

История. Заметки

Первые попытки построения ЭС для Военных целей (зенитные орудия, ТАУ) (1939-1945)

Системы автоматического Регулирования (САР) Владимира Викторовича Солодовникова (1939-1941)

“Кибернетика” Норберта Винера (1948)

Взлом Энигмы Марианом Реевским

Проект “Бомба” (Ежи Рожицкий, Генрих Зыгальский)

Передача чертежей в Кабацком лесу

Проект “Колосс” (Алан Тьюринг) (1935-41)

Машина Конрада Цузе (1941) Планкалкюль (1948)

История. Заметки

ENIAC
Джона Мокли
И
Джона Преспера Эккерта
(1943, 1946*)
*** публично представлен в Пенсильванском университете**

МИР
(Машина для инженерных Расчетов)
В.М. Глушкова
(1967)

Общегосударственная автоматизированная система учёта и обработки информации (ОГАС)
+
Единая государственная сеть вычислительных центров (ЕГСВЦ)
Анатолия Ивановича Китова
и
Виктора Михайловича Глушкова
(1958 – 1961* (1990))
*** статья в The Washington Post “Перфокарта управляет Кремлём”**

История. Заметки

ENIAC
Джона Мокли
И
Джона Преспера Эккерта
(1943, 1946*)
*** публично представлен в Пенсильванском университете**

МИР
(Машина для инженерных Расчетов)
В.М. Глушкова
(1967)

Общегосударственная автоматизированная система учёта и обработки информации (ОГАС)
+
Единая государственная сеть вычислительных центров (ЕГСВЦ)
Анатолия Ивановича Китова
и
Виктора Михайловича Глушкова
(1958 – 1964* (1990))
*** статья в The Washington Post “Перфокарта управляет Кремлём”**

История. Заметки

Метод опорных векторов
Support vector machine
Алексея Яковлевича Червоненкиса
и
Владимира Наумовича Вапника
(1963)

Метод “Случайный лес” (“Random Forest”)
Юрия Леонидовича Павлова
(1977)

Bootstrap
Бредли Эйфона
(1977)

Bagging (Bootstrap Aggregating)
Лео Бреймана
(1994)

Boosting
Роберта Шапира
(1990)

История. Заметки

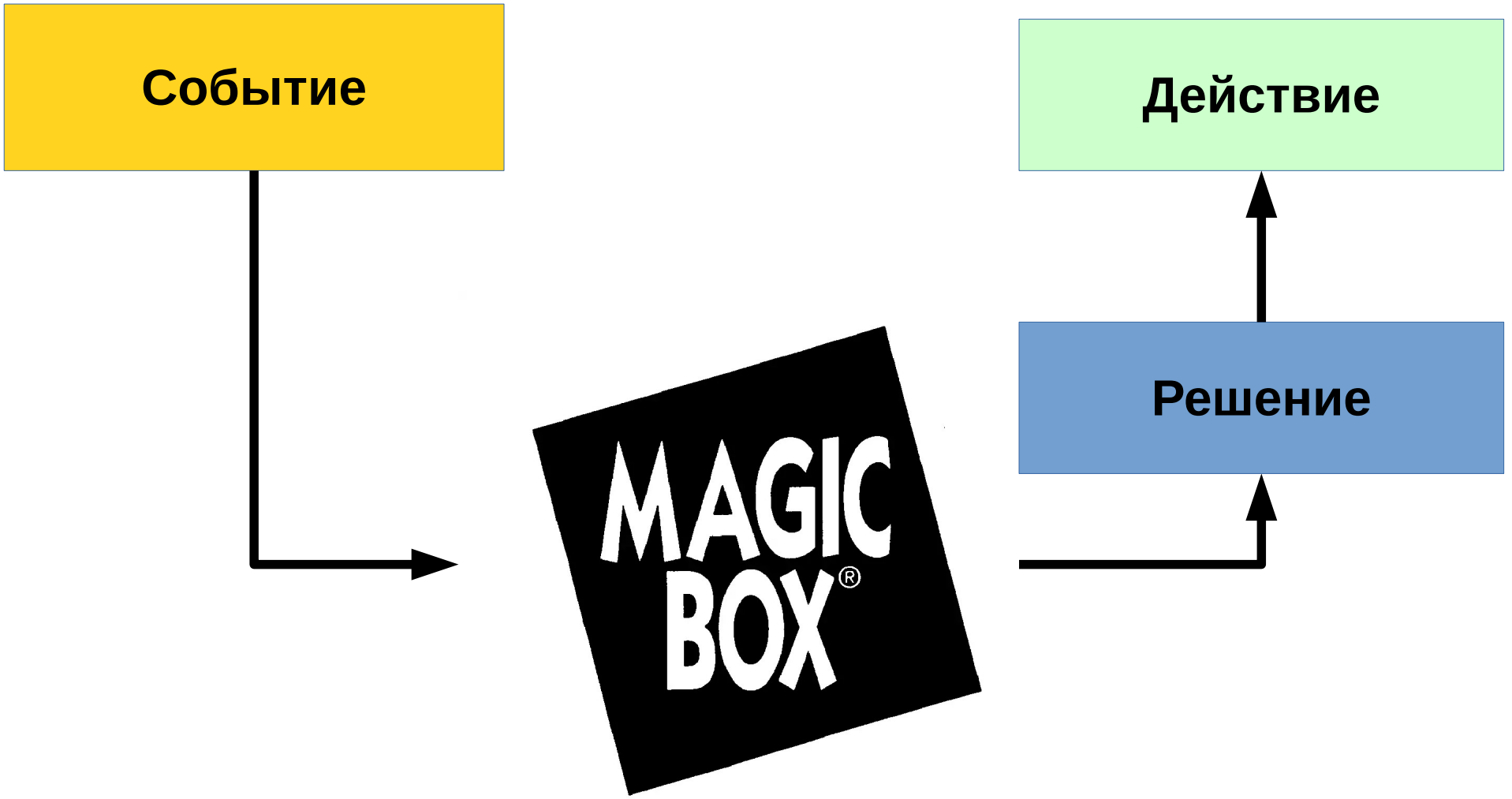
Байесовские сети
Джуды Перла
Judea Pearl
(1988)

“Закон Мура” и появление доступных персональных компьютеров открыло новую эпоху в ML и DM
(1990 – н.в.)

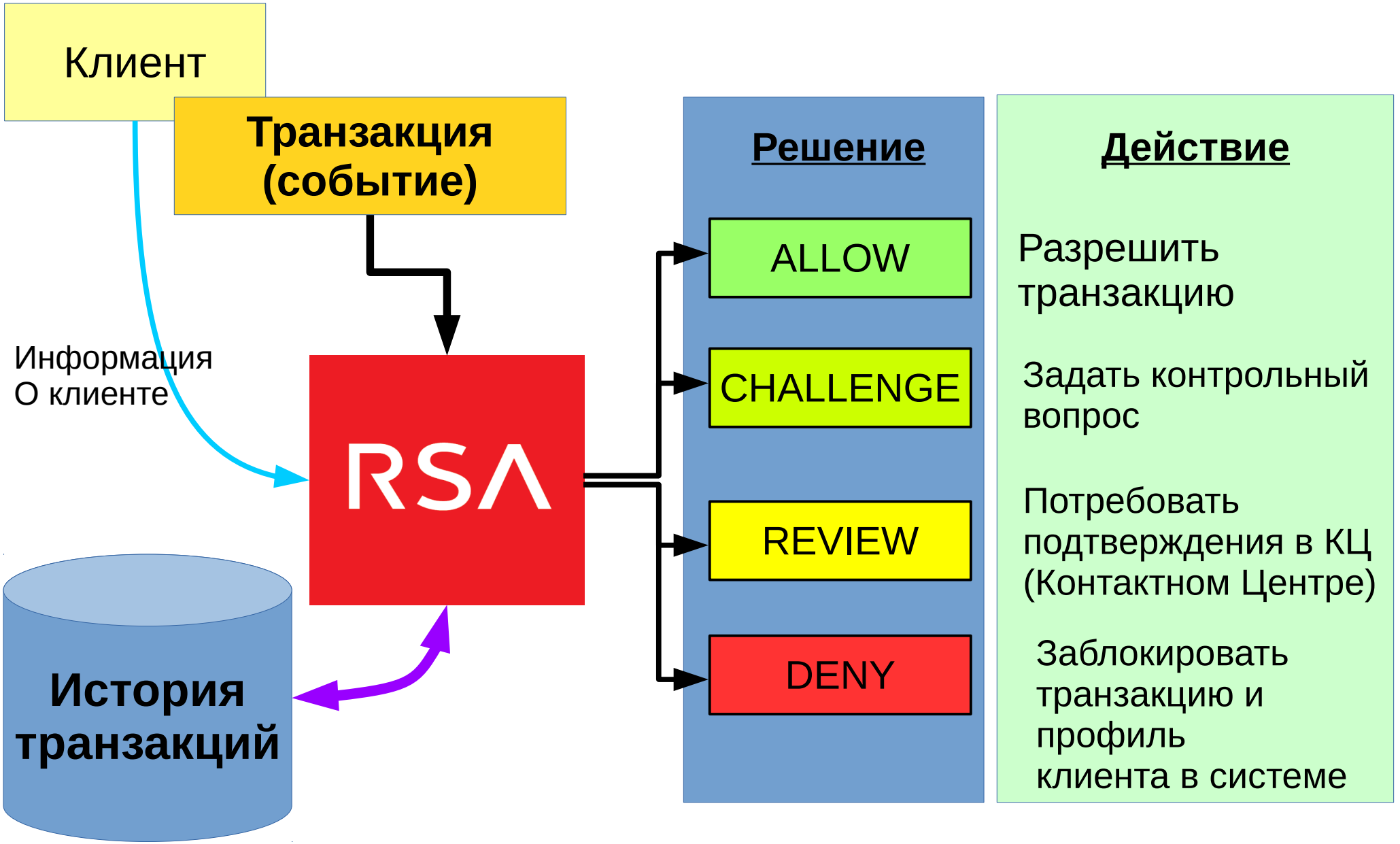
Журналист Клиффорд Линч
в журнале Nature
вводит понятие BigData
(2008)

первая* NoSQL СУБД
“Strozzi NoSQL”
Карло Стрози
(1998)
***после появления релационной алгебры**

ЭС сегодня



ФМ система RSA



Примеры задач ЭС для ИБ

- Банковский фрод (фрод-мониторинг, ФМ)
- Pay per click фрод
- Call фрод
- Обнаружение “социальной инженерии”

Примеры задач ЭС для ИБ

- Распознавание картинок, голоса, видео (ЦРТ)
- Распознавание жестов, мимики, рукописного ввода
- Распознование лиц
- Дактилоскопия

Примеры задач ЭС для ИБ

- Задача распознавания человек / робот
- Создание стратегий и тактик
- Поведенческий анализ botnet'ов.
- Обнаружение фазинга
- Обнаружение аномалий в системе

Примеры задач ЭС для ИБ

- Обнаружение спама
- Имитационное моделирование
- Crowdsourcing. Анализ качества введенной информации.
- Стеганография & стегоаналитика

Оценка качества ЭС

Как понять, хорошо работает **Magic Box** или плохо?

Ошибки 1-го и 2-го рода

Ошибка I рода	Ошибка II рода
α - ошибка	β - ошибка
false positive (fp)	false negative (fn)
ложная сработка	**пропуск цели**

Ложная сработка

VS

Пропуск цели

Что важнее?

Примеры

- ЭС диагностирования раковой опухоли.

  Ложная сработка – это когда ЭС говорит здоровому пациенту, что у него обнаружен рак

  Пропуск цели – система говорит больному раком пациенту, что он здоров.

- ЭС обнаружения превышения скорости автомобиля

  Ложная сработка – штраф будет выписан добропорядочному водителю.

  Пропуск цели – лихач не будет платить штрафа

Сложные случаи

- Фрод мониторинг

  Ложная сработка – заблокировать легитимную транзакцию

  Пропуск цели – позволить хакерам украсть деньги клиента

Вынесение решений системой

- **Булевое решение**: либо 0, либо 1. Как в суде: виновен – не виновен
- **Вероятностное решение**: $p \in [0, 1]$

  Показывает, какова вероятность. 0 – однозначно нет, 1 – однозначно да. 0.5 – неопределено.

Сведение вероятностного решения к булевому

- p > 0.5 – виновен
- p <= 0.5 – не виновен
- Вероятностное решение всегда сводится к булевому
- **Замечание**: p – это ***априорная*** вероятность. Если система “плохая”, то p может вообще не иметь связи с реальностью.

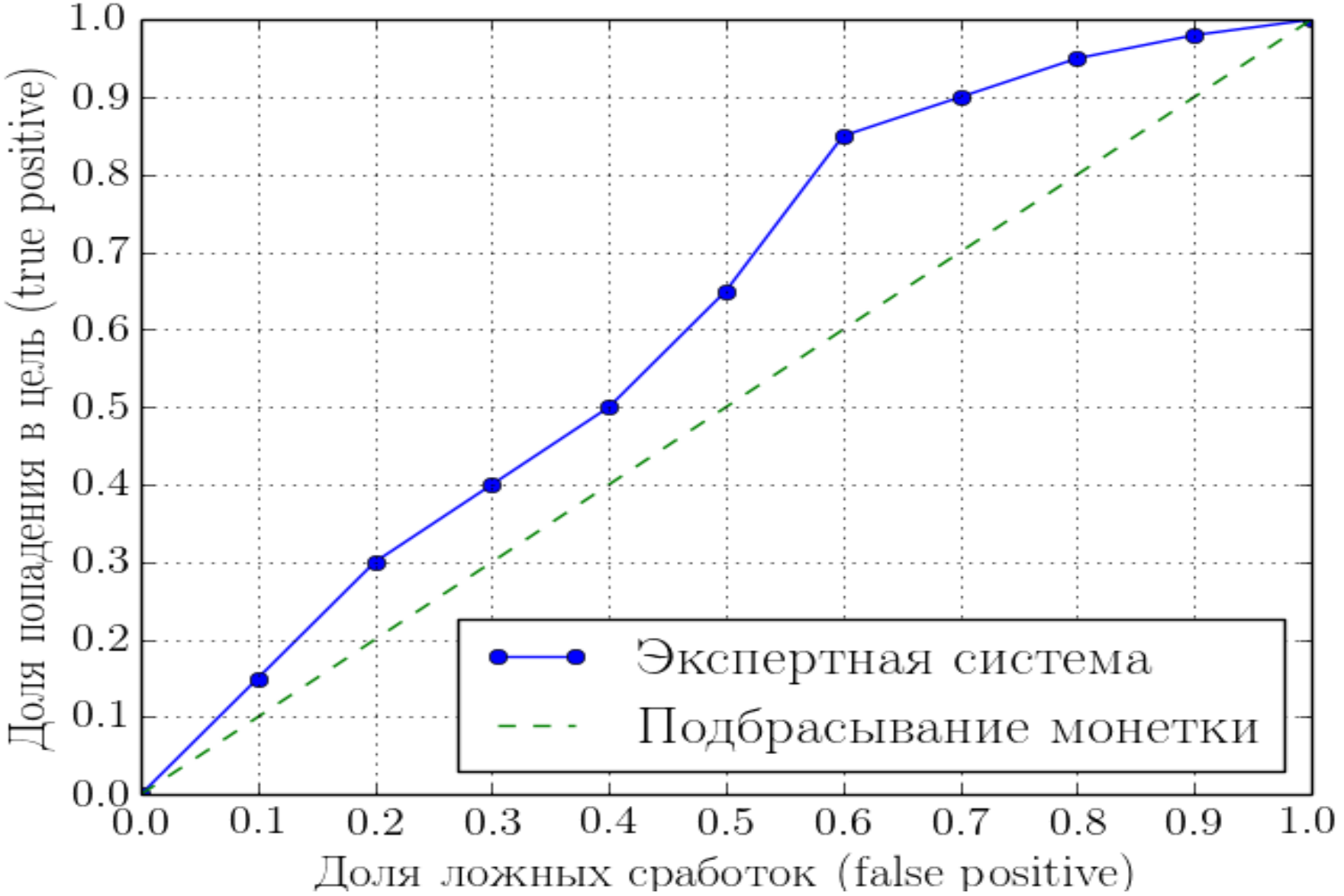
Как повлиять на вероятность **ложной сработки**?

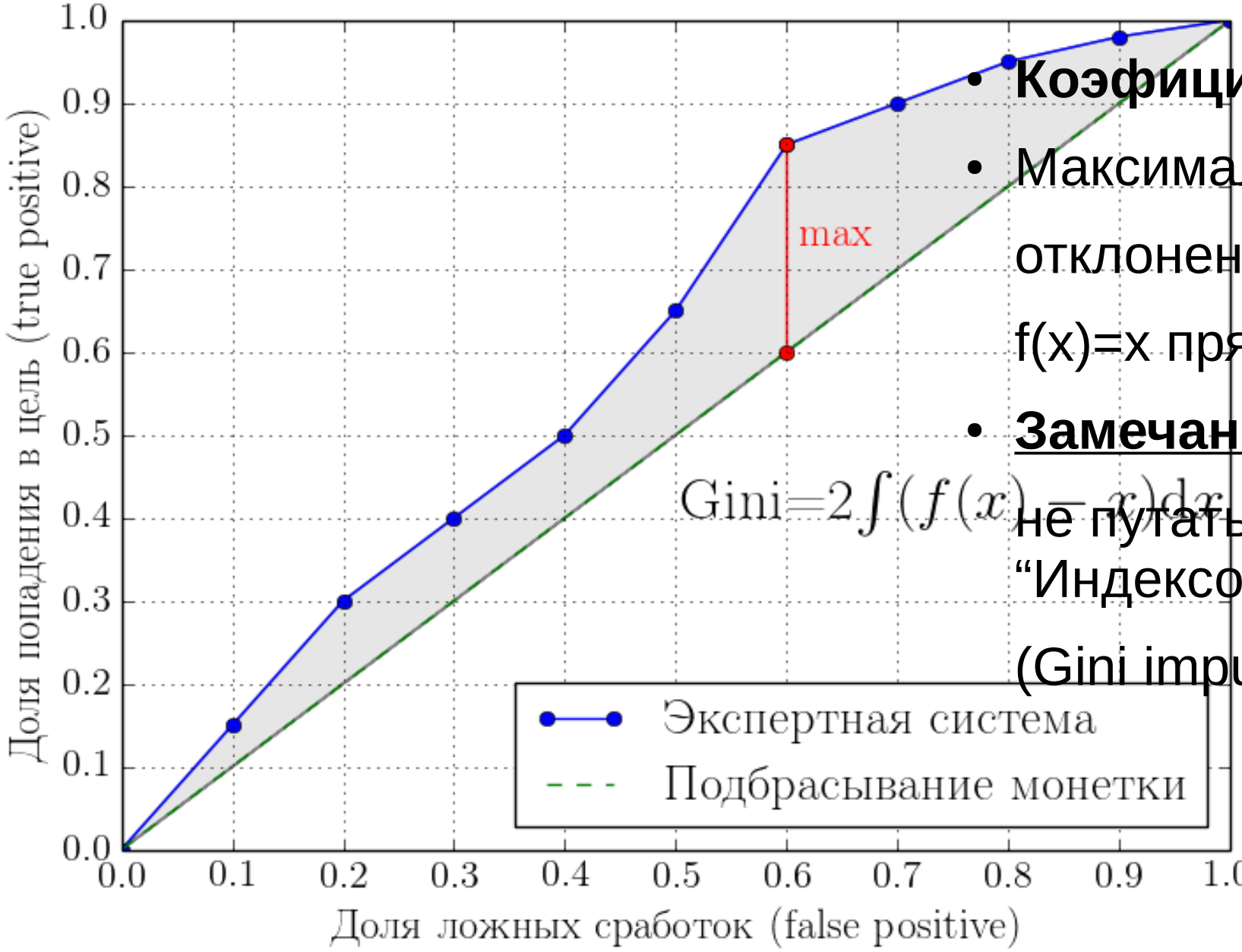
Как повлиять на вероятность **пропуска цели**?

Изменение порога решающего правила

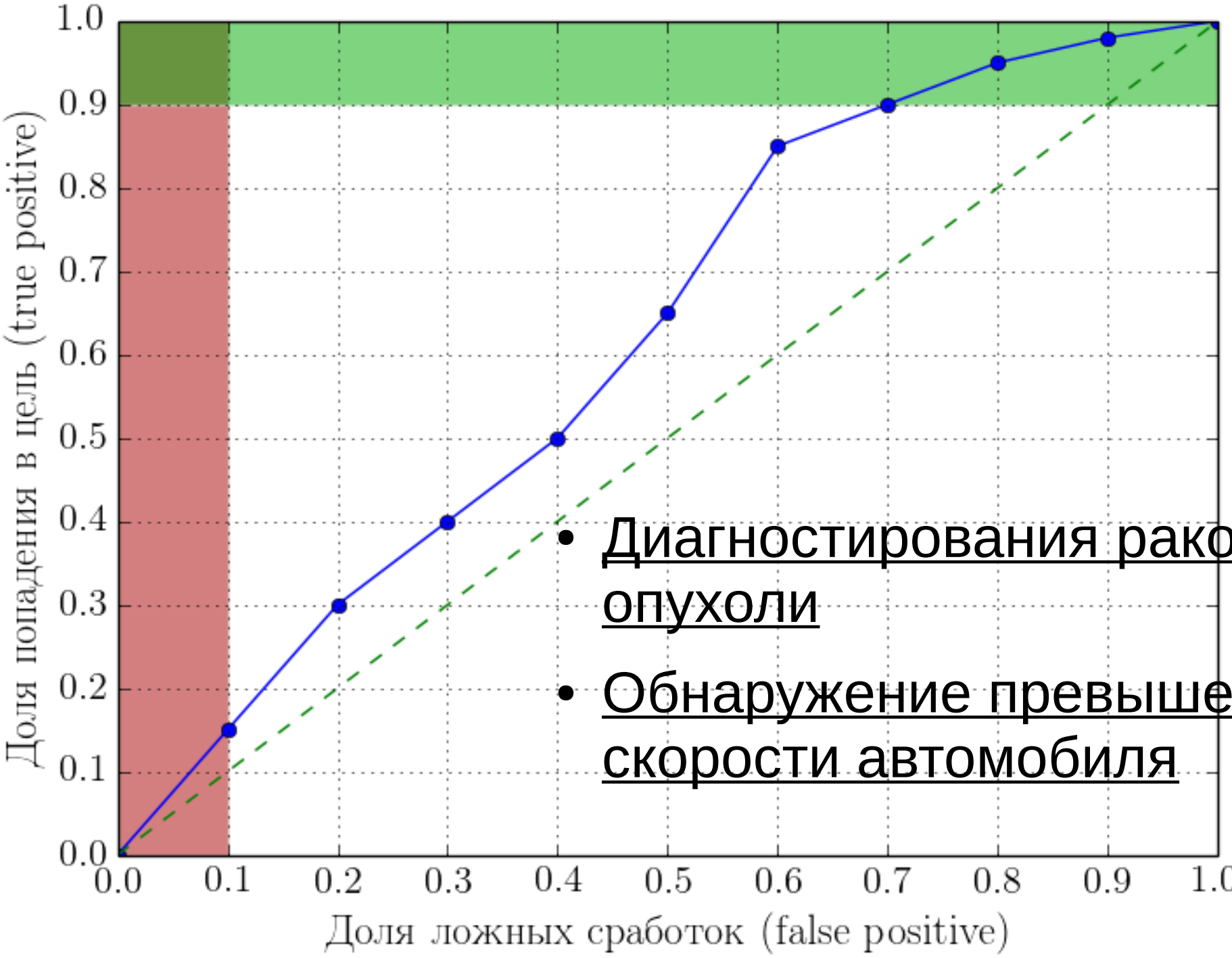
- Было: p > 0.5 – виновен; p <= 0.5 – не виновен
- Поменяли отсечку: p > 0.8 – виновен; p <= 0.8 – не виновен
- Что стало с ***апостериорными*** вероятностями ложной сработки & пропуска цели?
- Что не уменьшилось, что не увеличилось?

ROC-кривая (receiver operating characteristic)

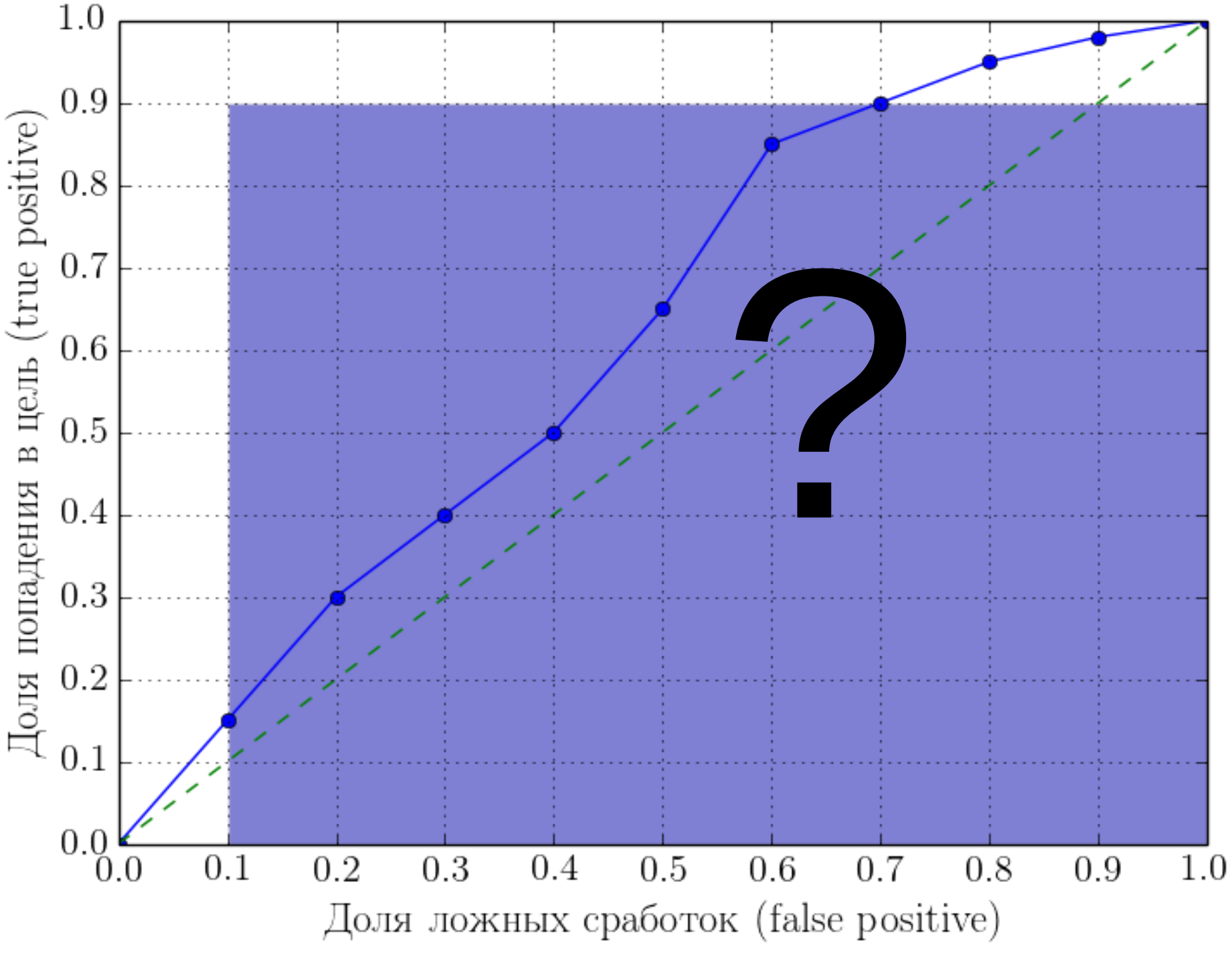




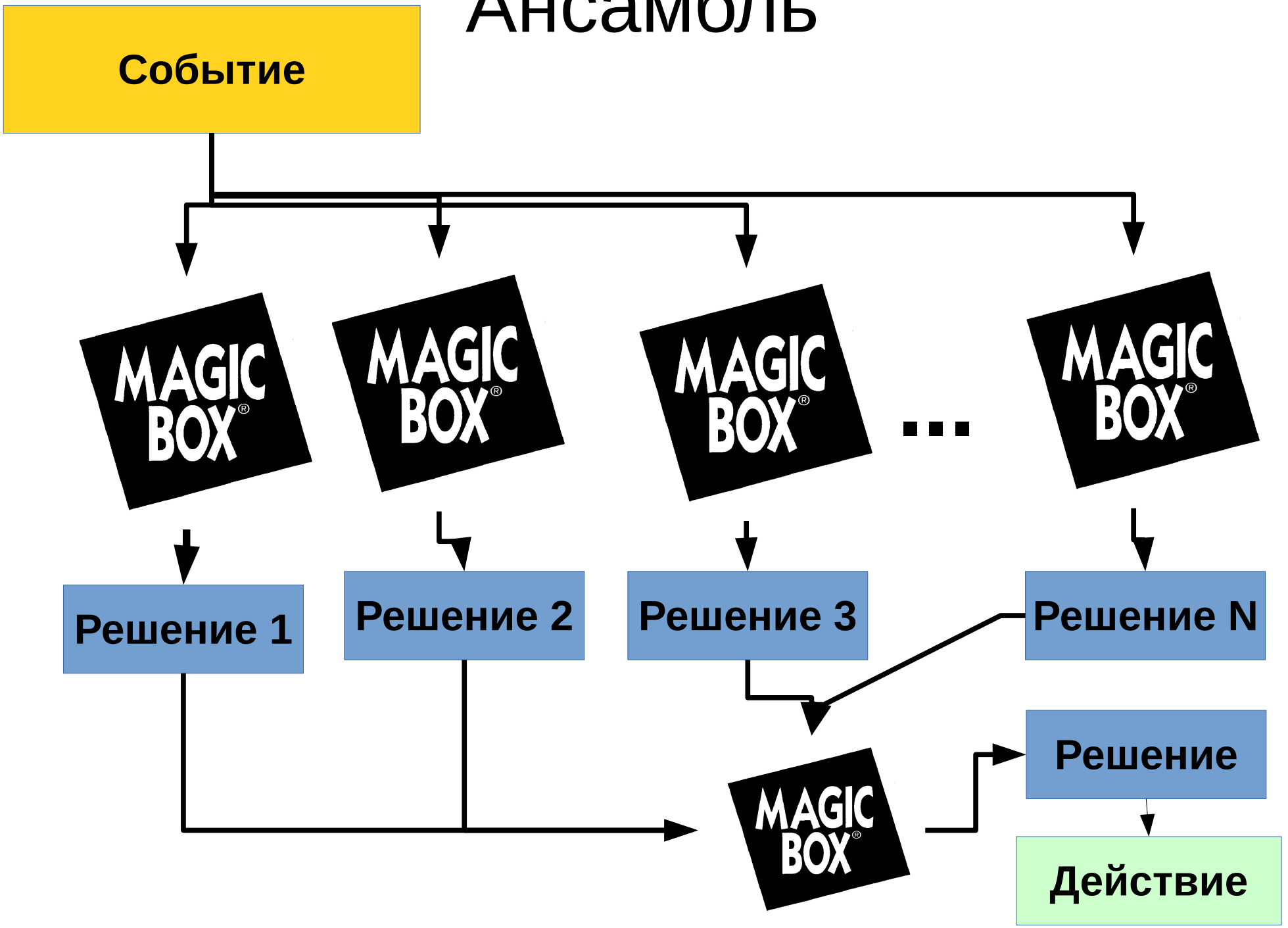
- **Коэффициент** Джини
- Максимальное отклонение от f(x)=x прямой
- **<u>Замечание</u>**: не путать с “Индексом Джини” (Gini impurity)



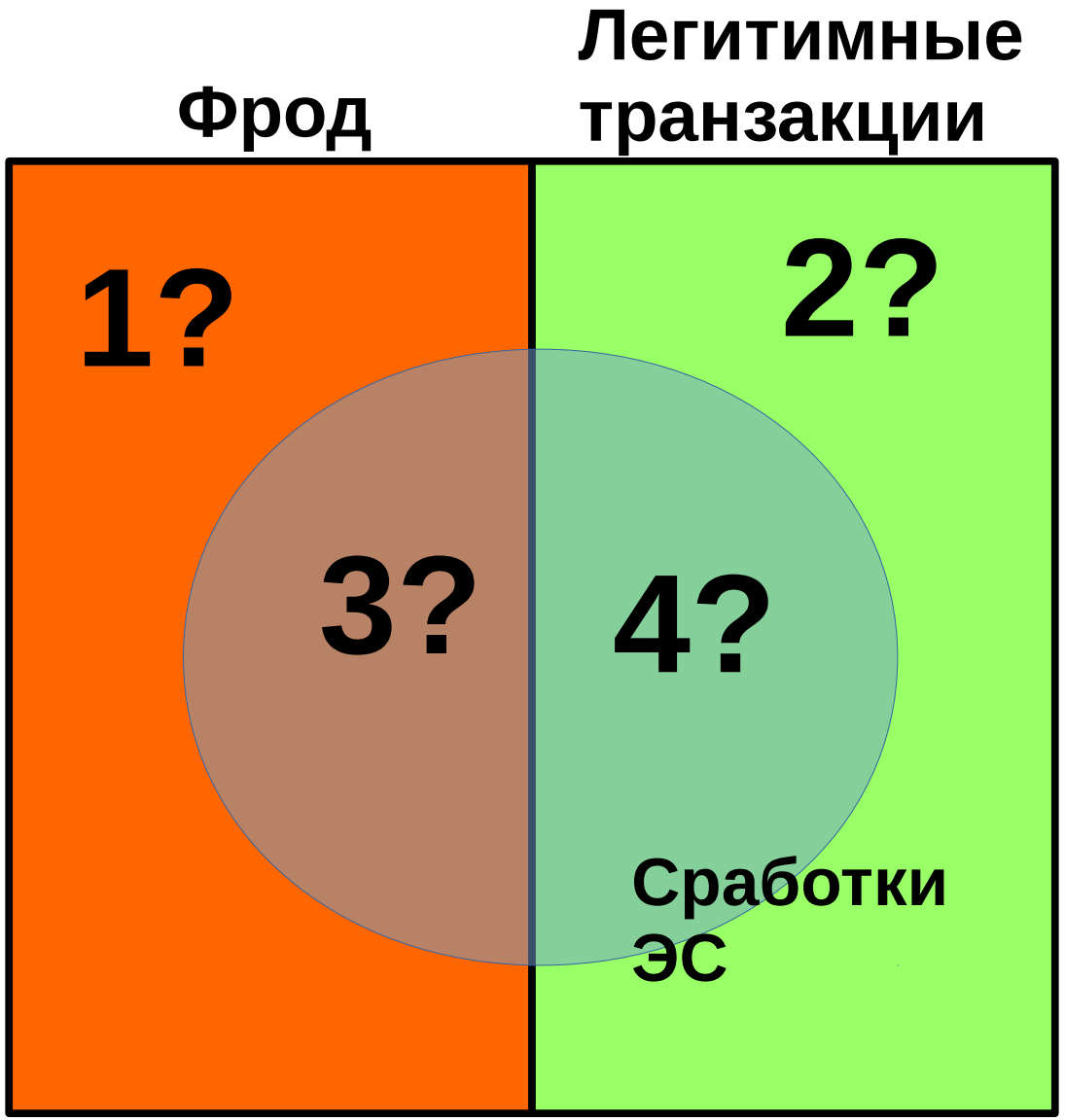
- Диагностирования раковой опухоли
- Обнаружение превышения скорости автомобиля


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Доля попадения в цель (true positive)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Доля ложных сработок (false positive)
?

Ансамбль


Событие
MAGIC BOX®
MAGIC BOX®
MAGIC BOX®
...
MAGIC BOX®
Решение 1
Решение 2
Решение 3
Решение N
MAGIC BOX®
Решение
Действие

tp, fp, tn, fn.
Полнота, точность, чувствительность.



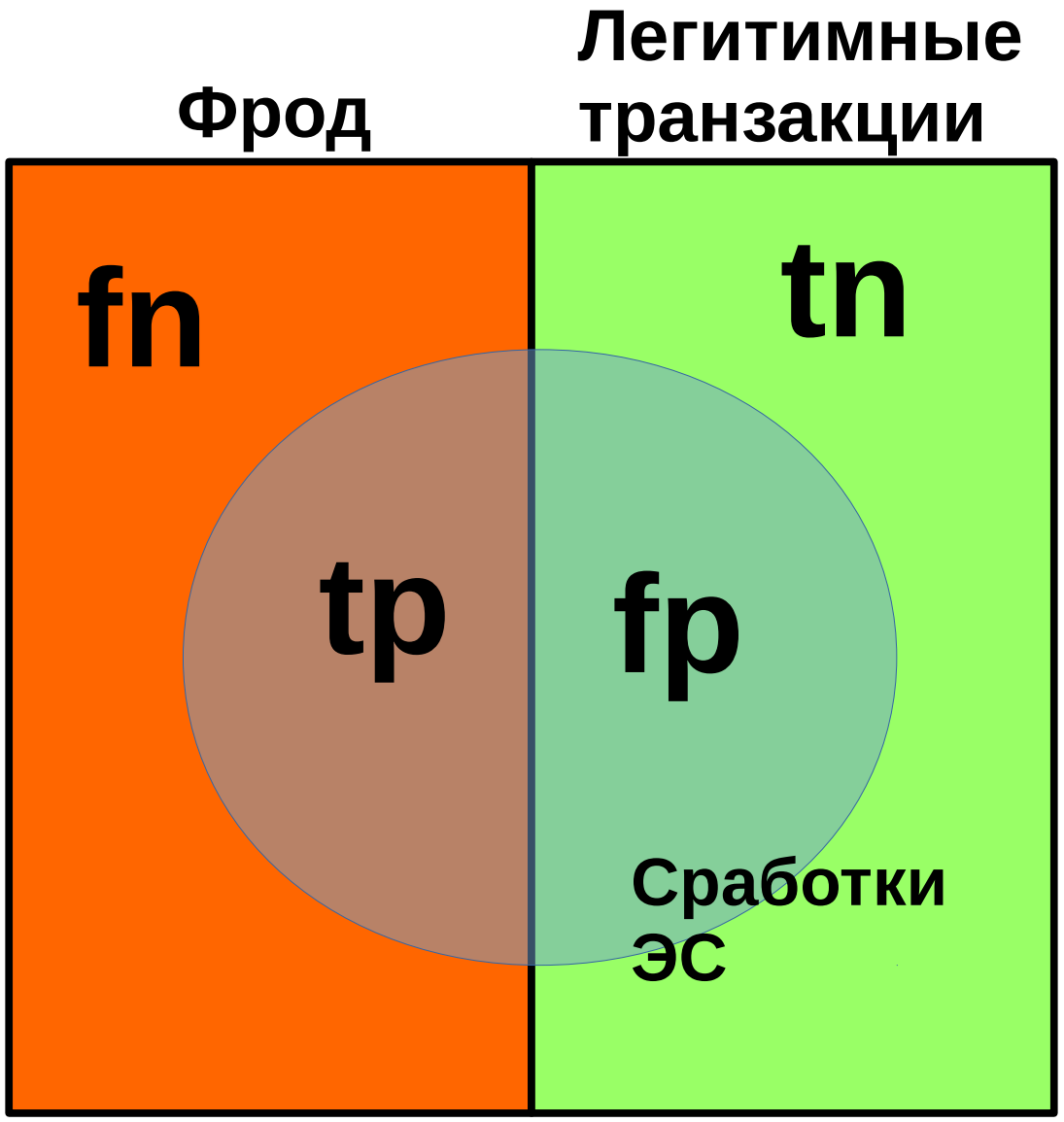
Где здесь ???

True positive

True negative

False Positive

False Negative



Полнота

$$Recall = \frac{tp}{(tp + fn)}$$

Точность

$$Precision = \frac{tp}{(tp + fp)}$$

“Аккуратность”, “Тщательность”

$$Accuracy = ACC = \frac{tp + tn}{tp + tn + fp + fn}$$

F-мера

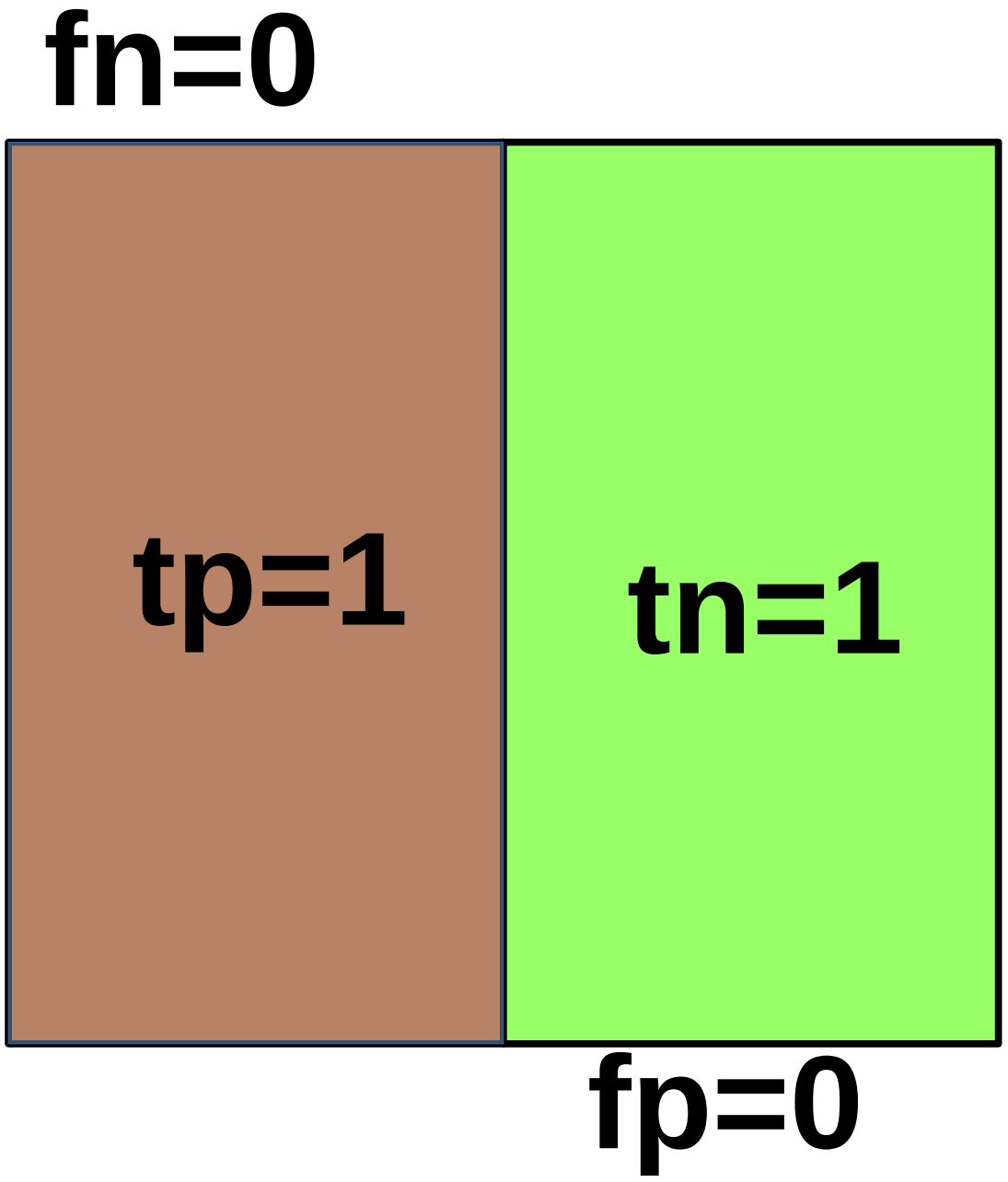
- Единый критерий,зависящий от precision и recall

$$F_\beta = \frac{(1+\beta^2)\cdot precision \cdot recall}{\beta^2 \cdot precision + recall}$$

- В частности

$$F_1 = \frac{2 \cdot precision \cdot recall}{precision + recall}$$

ЭС “Бога”



Идеальный вариант

Вопрос на засыпку:
Чему равен
коэфициент Джини?

System Stability Index

- Задается отношение двух величин на момент времени **t**: **x**, и **y=f(x, t).**

  Например **y** – точность (или полнота) ЭС в зависимости от **x=tp**.

- Выбирается **n**.
- Выбираются значения xi: {x1, x2, … xn}.
- Определим **q >= 2** (обычно **q=2** или **q=e**)

$$SSI(t_1,t_2)=\sum_{i=\overline{1,n}}\left(f(x_i;t_1)-f(x_i;t_2)\right)\cdot\log_q\left(\frac{f(x_i;t_1)}{f(x_i;t_2)}\right)$$

Cost-Benefit Analysis

- Все свойства системы оцениваем (в рублях, в попугаях, в условных баллах)
- Определяются затраты и им так же дается оценка.
- Составляется СЛАУ
- Определяем параметры системы, дающие максимальные СВА
- В некоторых случаях строго решить СЛАУ невозможно (или невероятно трудно) и создают ЭС для подсчета СВА другой ЭС ;))

Value at risk (VaR)

- Определяются риски.
- Для каждого риска вводиться вероятность риска и ущерб риска
- Подсчитывается цена риска. Обычно:

$$risk_i(cost_i, p_i) = cost_i \cdot p_i$$

- Определяются параметры, влияющие на риски.
- Составляется СЛАУ
- Определяются параметры, дающие минимальный VaR

Требования к ***конструктору*** ЭС

- Хороший навык **программирования**
- Умение работать с **СУБД**. Обычно:
  - Csv ;)
  - MariaDB (MySQL)
  - Oracle
  - Postgre
  - MongoDB
  - Redis
  - мир Hadoop,
  - … …

Требования к ***конструктору*** ЭС

- **Математика**:
  - Теор.вер
  - Мат.статистика
  - Теория алгоритмов
  - Комбинаторика
  - Аналитическая геометрия
  - Теория графов
  - Дискретная математика
- **Предметная область!**

  (в нашем случае – ИБ)

Резюме

- Для оценки качества ЭС ***как “продукта”*** не требуется знать, как ЭС устроены.
- Для того, чтобы понять, как ***улучшить*** ЭС (и можно ли улучшить) – нужно знать, как ЭС устроена
- Основные навыки и умения, необходимые конструктору ЭС:
  - Предметная область (в нашем случае ИБ)
  - Хороший программист
  - Хороший математик
  - Владение СУБД