

Лекция

27–28 апреля

Разрешимость и теорема Райса



Разрешимость и перечислимость

Мы должны знать. Мы будем знать.

— Давид Гильберт

Разрешитель против распознавателя

Определение 1 (Два класса машин)

- **Разрешитель** всегда останавливается и отвечает да или нет.
- **Распознаватель** принимает слова своего языка, а вне его может заикнуться.

Язык разрешим — есть разрешитель, язык перечислим — есть распознаватель.

Разрешимость требует двух правильных ответов, распознаваемость — только одного.

Примеры: разрешимо и распознаваемо

Пример (Два языка)

- $a^n b^n$ разрешим: машина вычёркивает пары и всегда останавливается.
- HALT распознаваема: универсальная машина симулирует M на w , но закливание симуляции не даёт ответа «нет».

Вычислимо — не значит практически решаемо.

Разрешимая задача может требовать астрономического времени: граница «в принципе» и граница «на практике» — разные границы.

Перечислимость

Определение 2 (Перечислимость)

Язык L **перечислим**, если есть МТ, которая печатает все элементы L — каждый хотя бы раз и ничего кроме.

Эквивалентно: есть распознаватель L .

Распознаватель и перечислитель — две маски одной силы: запустить машину на всех входах параллельно и печатать принятые.

Перечислимые языки — то, что машина способна хоть как-то предъявить.

С двух сторон

Теорема 1 (Перечислимость плюс ко-перечислимость)

L разрешим тогда и только тогда, когда перечислимы L и $\bar{L}$.

Доказательство

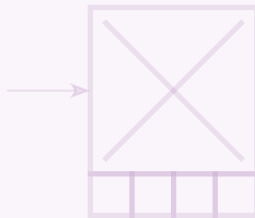
Докажем в обе стороны.

Если L разрешим, оба распознавателя получают готовые ответы разрешителя.

Обратно: пусть есть распознаватели M для L и N для $\bar{L}$.

Запустим их параллельно, чередуя шаги: свой язык имеет ровно один из них, он и остановится — его приём и есть ответ.

Дополнение HALT даже не перечислимо — не всё честно сформулированное доступно машине хотя бы наполовину.



Теорема Райса

Остерегайтесь ошибок в приведённом выше коде: я лишь доказал его корректность, но не исполнил его.

— Дональд Кнут

Свойство языка или свойство машины?

Свойство P описаний машин называется экстенциональным, если оно зависит только от языка:

$$L(M_1) = L(M_2) \Rightarrow (\langle M_1 \rangle \in P \leftrightarrow \langle M_2 \rangle \in P).$$

Свойство нетривиально, если есть машины обоих сортов: одна с P , другая без.

«Язык пуст» — экстенционально, «не более 100 состояний» — свойство текста, а не языка.

Всякий вопрос «что программа вычисляет?» экстенционален, всякий «как она устроена?» — нет.

Теорема Райса

Теорема 2 (Теорема Райса)

Всякое нетривиальное экстенциональное свойство распознаваемых языков неразрешимо.

Иначе: язык $\{\langle M \rangle \mid L(M) \text{ удовлетворяет } P\}$ неразрешим для каждого нетривиального P .

Единственные разрешимые свойства — тривиальные: выполненные для всех языков или ни для одного.

После Райса вопрос «неразрешимо ли?» заменяется вопросом «тривиально ли?».

Идея доказательства

Доказательство (Набросок)

Докажем сведением HALT к P .

Без ограничения общности $\emptyset$ не удовлетворяет P — иначе рассуждаем про дополнение свойства.

По нетривиальности есть машина A с $L(A) \in P$.

Строим M' : на любом входе x она симулирует M на w , а при остановке симуляции работает как A .

M остановилась на w — значит, $L(M') = L(A)$ и свойство есть. M заиклилась — значит, $L(M') = \emptyset$ и свойства нет.

Разрешитель P дал бы разрешитель HALT — противоречие.

Что закрыто Райсом?

Пример (Неразрешимые свойства языков)

- Пуст ли $L(M)$? Конечен ли он?
- Регулярен ли $L(M)$? Контекстно-свободен?
- Верно ли $L(M) = \Sigma^*$? Содержит ли $L(M)$ слово 01?

Все — нетривиальные свойства языка — все неразрешимы одной теоремой.

Граница: «останавливается ли M на ε » — свойство поведения, а не языка, Райсу не подчиняется.

Но и оно неразрешимо — отдельным сведением из прошлой лекции.

Инженерный итог

Статический анализ, антивирусы, верификаторы проверяют экстенсionalesные свойства — и обречены на приближения.

Точность и разрешимость несовместимы: инструмент либо молчит на части программ, либо врёт.

Поэтому анализаторы выдают «да», «нет» и «не знаю» — третий ответ обязан существовать.

За пределами машин Тьюринга

*Любая достаточно развитая технология неотличима от
магии.*

— Артур Кларк

Оракулы

Определение 3 (МТ с оракулом)

МТ с дополнительной лентой-оракулом: записав строку и перейдя в состояние запроса, машина за один шаг узнаёт, принадлежит ли строка языку A .

Оракул — чёрный ящик, решающий A : подпрограмма, отвечающая мгновенно.

Пример (Оракул для HALT)

Даже он не решает тотальность: вопрос «для всех ли x машина останавливается?» требует бесконечной серии запросов — по одному на каждый вход.

Лестница иерархии

Оракул для HALT открывает часть задач — и порождает новую проблему остановки: уже для машин с оракулом.

Тот же шаг повторяется наверху — иерархия не кончается.

Кванторная глубина — мост к логике первого порядка из первого семестра:

- $\exists t$: «остановится за t шагов» — перечислимость
- $\forall t$: «не остановится никогда» — ко-перечислимость
- $\forall x \exists t$ — следующий уровень, тотальность.

Гипервычисления

Гипервычисления — спекуляции о машинах, выходящих за тезис: бесконечное ускорение, оракулы в физике, точные действительные числа за один шаг.

Тезис Чёрча–Тьюринга — не теорема, и физика не обязана подчиняться определению.

Но за девяносто лет ни одна реализация не вышла за класс тьюринг-вычислимого: гипотез хватает, работающих машин — нет.

Что с этим делать инженеру?

Если задача неразрешима на МТ, её не решит ни один язык программирования.

Значит: искать разрешимые фрагменты, ограничивать класс входов, строить приближения.

Умение распознать неразрешимость — экономит месяцы охоты за несуществующим алгоритмом.

Лекция 12: что мы поняли?

- Разрешитель против распознавателя: разрешимость требует ответа на каждом входе, перечислимость — только предъявления своих слов.
- Перечислимость с двух сторон даёт разрешимость. Дополнение HALT показывает, что даже перечислимости может не быть.
- Теорема Райса: всякое нетривиальное свойство языка машины неразрешимо — доказательство сводит остановку, подменяя язык машины по ответу решателя.
- Выход за пределы машин Тьюринга — оракулы и лестница иерархии: каждая ступень открывает задачи и рождает свою неразрешимость.
- Тезис Чёрча–Тьюринга остаётся определением, проверяемым опытом, — гипервычисления пока не вышли из спекуляций.

Вопросы для самопроверки

1. Чем разрешитель отличается от распознавателя и какой класс языков у каждого?
2. Почему HALT перечислима, а её дополнение — нет?
3. Сформулируйте и докажите: перечислимость L и $\bar{L}$ даёт разрешимость.
4. Сформулируйте теорему Райса и проверьте, применима ли она к свойству «у машины ровно три состояния».
5. Воспроизведите идею доказательства Райса: как строится машина M' ?
6. Почему оракул для HALT не решает тотальность?
7. Что означает «тезис — не теорема» и что могло бы его пошатнуть?

Чтение к следующей паре

- m26 — «Теорема Райса», «За пределами машин Тьюринга».
- Флеш-опросник — в начале лекции 13.
- Контрольная 3 (машина Тьюринга и разрешимость) — на отдельной паре на следующей неделе.