

EUCLID'S ALGO (OCD)

- if $y = 0$ return x
- if $y \neq 0$ return $\text{Euclid}(y, x \bmod y)$

POTENTIAL (LEAKY BUCKET):

- Unit of time
- How to measure start $S_0 \geq 0$
- Show $S_{t+1} < S_t$

** find good potential function **
SIZE OF INPUT = $O(\log x)$

MASTER THEOREM:

$$T(n) = k T\left(\frac{n}{b}\right) + O(n^d)$$

k : # of subproblems

b : size of subproblem

n^d : time to "merge"

$$T(n) = \begin{cases} O(n^d) & k < b^d \\ O(n^d \log n) & k = b^d \\ O(n^{\log_b k}) & k > b^d \end{cases}$$

$$T(n) = k T\left(\frac{n}{b}\right) + O(n^d \log^w n)$$

$$T(n) = \begin{cases} O(n^d \log^w n) & k/b^d < 1 \\ O(n^d \log^{w+1} n) & k/b^d = 1 \\ O(n^{\log_b k}) & k/b^d > 1 \end{cases}$$

$$f(n) = O(g(n)) \rightarrow |f(n)| \leq M|g(n)|, \forall n > n_0$$

$$f(n) = \Omega(g(n)) \rightarrow |f(n)| \geq M|g(n)|, \forall n > n_0$$

$$f(n) = \Theta(g(n)) \rightarrow f(n) = O(g(n)) \wedge f(n) = \Omega(g(n))$$

$$f(n) = o(g(n)) \rightarrow |f(n)| < M|g(n)|, \forall n > n_0$$

$$f(n) = w(g(n)) \rightarrow |f(n)| > M|g(n)|, \forall n > n_0$$

Proving $f(n) = O(g(n))$:

- MASTER THM
- $\lim_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| < c$ for $c \in \mathbb{R}$

$$(L'Hopital: \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{f'(n)}{g'(n)})$$

$$\text{Note: } \frac{d}{dx} ax \rightarrow \ln(a)x^a$$

Find witnesses M, n_0

Alphabet Σ : non-empty, finite set of symbols

Σ^n : $\Sigma \times \Sigma \times \dots \times \Sigma$, set of all strings of length n from Σ

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots$

Language L : $L \subseteq \Sigma^*$

DFA: $(Q, \Sigma, \delta, q_0, F)$

$\rightarrow Q$: set of states

$\rightarrow \Sigma$: alphabet

$\rightarrow \delta: Q \times \Sigma \rightarrow Q$: transition function

$\rightarrow q_0 \in Q$: initial state

$\rightarrow F \subseteq Q$: accepting state(s)

DFA M "accepts" x iff ends in a final state $\in F$ after reading x

M decides regular language

$L(M) = \{x \in \Sigma^* \mid M \text{ accepts } x\}$

If L_1, L_2 one regular languages

$\Sigma^* - L(\bar{L}) = L^*$

L^R (reverse) — flip direction?

$L_1 \cup L_2, L_1 \cap L_2$

are all regular languages

No DFA decides $\{0^n 1^n \mid n \geq 0\}$

It is possible to decide if 2 DFAs decide same language

CHURCH-TURING THESIS

Every problem computable by mechanical device is computable by a Turing Machine

Two models of computation equivalent if one can simulate the other (prove both ways)

REG EXES:
*: 0 or more
+: 1 or more
 Δ : all uppercase letters

CLOSEST PAIR (2D):

- Set points by x-coord
- Find closest pair in each half
- $\Delta: \min(S_1, S_2)$
- To find close red/blue pair: slide $\Delta \times 2$ rectangle down (for each P E filtered in 6-step compare to 7 lower points)

$$\gg T(n) = 2T(n/2) + O(n) = O(n \log n)$$

KARASTRUBA MULTIPLICATION:

$$\gg x = a \cdot 10^{n/2} + b; y = c \cdot 10^{n/2} + d$$

$$\gg xy = ac \cdot 10^n + (ad+bc) \cdot 10^{n/2} + bd$$

$$\gg ac; bd; ad+bc = (a+b)(c+d) - ac - bd$$

$$\gg T(n) = 3T(n/2) + O(n) \rightarrow O(n^{\log_2 3})$$

\gg Do 3 muls instead of 4!

LONGEST INCREASING SUBSEQUENCE

- Sequence of numbers, $s_i < s_{i+1}$
- $LIS(\Delta) = LIS$ of array of nums Δ
- Can't just directly compute — solve a related subproblem
 - $L(\Delta) = LIS$ that includes last num

$$\gg L(\Delta) = |\Delta|$$

$$\gg x = \max_{i \in \Delta} \{ \Delta[i] \}$$

\gg Find all prev entries that are smaller, take max of those $L(\Delta)$, add 1

BIG-O DEFINITIONS + PROPS

$$f(n) = O(g(n)) \rightarrow |f(n)| \leq M|g(n)|, \forall n > n_0$$

$$f(n) = \Omega(g(n)) \rightarrow |f(n)| \geq M|g(n)|, \forall n > n_0$$

$$f(n) = \Theta(g(n)) \rightarrow f(n) = O(g(n)) \wedge f(n) = \Omega(g(n))$$

$$f(n) = o(g(n)) \rightarrow |f(n)| < M|g(n)|$$

$$f(n) = w(g(n)) \rightarrow |f(n)| > M|g(n)|$$

PROVE $f(n) \approx O(g(n))$:

Use Master Thm to convert

$$\lim_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| = \infty \text{ (disregarding not enough)}$$

$$\gg \text{Show } g(n) = O(f(n)) \Rightarrow \lim_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| = 0$$

Argue 203-style: $\forall M, n_0, \exists n > n_0$ s.t. $|f(n)| > M|g(n)|$ (contradict n value)

FIND WITNESSES M, n_0

CHOMSKY'S HIERARCHY

Regular Languages \leftrightarrow Finite Automata

Context-Free Ls \leftrightarrow Pushdown Automata

Context-Sensitive Ls \leftrightarrow Linear-Bounded Aut.

Recursively-enumerable \leftrightarrow TURING (recognizable) Ls

MACHINES

TURING MACHINES

READ-ONLY TM \in DFA (Rabin + Scott)

CONSIST OF:

\rightarrow 2D infinite tape of cells — initially contains input, then blanks (λ)

\rightarrow tape head: can be moved L/R initially @ leftmost cell

reads symbol under head and writes symbol under head and rejects from tape alphabet $T \subseteq \Sigma \cup \{\lambda\}$

\rightarrow controlled by state machine — accepts/rejects on special states

$\bullet M = (Q, \Sigma, \delta, q_0, F)$

Q: set of states T : tape alphabet

Σ : input alphabet (Σ, λ , and more Σ)

\bullet $\Sigma^* = \Sigma \times \Sigma \times \dots \times \Sigma$

\bullet M accepts $x \iff$ ends @ accept

$L(M) = \{x \in \Sigma^* \mid M \text{ accepts } x\}$

M recognizes $L(M)$

\bullet M also DECIDES $L(M)$ if it halts on all inputs; \therefore it accepts all $x \in L(M)$ and rejects $x \notin L(M)$

\bullet CH (1978): No set S exists such that $|N| < |S| < |I|$

Paul Cohen (1963): CH independent of classical set theory axioms

DYNAMIC PROGRAMMING:

① Top - Down Recursive (inefficient if overlapping subproblems; space-efficient)

② Top - Down Memoization (faster time, more space, only computes needed subps)

③ Bottom - Up Table (faster time, more space, may calculate unnecessary subproblems)

LONGEST COMMON SUBSEQUENCE:

LCS($[1 \dots m]$, $[1 \dots n]$):

\gg If $m=0, n=0$: return 0

\gg If $x[m] = y[n]$: return $g + LCS([1 \dots m-1], [1 \dots n-1])$

\gg Else: $LCS([1 \dots m-1], [1 \dots n])$

\gg (Do using table)

RECURRANCE RELATIONS:

Rod - Cut

Given rod of length n , list P of prices

where $P[i]$ is price of length i ; find max profit of cutting rod + selling pieces

$\gg R(0) = 0; R(l) = P[l]$

$\gg R(n) = \max_{k \in \{0, \dots, n-1\}} \{P[k] + R(n-k)\}$

EDIT - DISTANCE:

\rightarrow Transform string A into B via substituting, deleting, inserting characters (cost s, d, i)

\rightarrow Find minimum-cost of transforming A to B

\gg Let $A = [1 \dots m]$, $B = [1 \dots n]$

\gg If $A[m] = B[n]$: $\rightarrow ED([1 \dots m-1], [1 \dots n-1])$

\gg Else: $\min_{i \in C: s_i > 0} \{ ED([1 \dots m-i], [1 \dots n-1]) \}$

\gg $ED([1 \dots m], [1 \dots n]) = \min_{i \in C: s_i > 0} \{ ED([1 \dots m-i], [1 \dots n-1]) \}$

\gg $LPS([1 \dots m], [1 \dots n]) = \min_{i \in C: s_i > 0} \{ LPS([1 \dots m-i], [1 \dots n-1]) \}$

\gg $LPS([1 \dots m], [1 \dots n]) = \max_{i \in C: s_i > 0} \{ LPS([1 \dots m-i], [1 \dots n-1]) \}$

PROVING NOT ALL BINARY LANGUAGES DECIDABLE

Given ALL be set of all binary languages; D be decidable binary languages

Show that $|D| = |N| < |ALL| = |\mathbb{R}|$ (countably vs. uncountably ∞)

RECALL: Δ is set countable if bijection to N (can list in some order)

Σ, \mathbb{Z}^* is countable — just list LEXICOGRAPHICALLY (length, then alphabetical)

Similarly, all valid C++ programs countable (lexicographical ordering)

As C++ programs \equiv Turing Machines \leftrightarrow Languages, $|D|$ is thus COUNTABLE.

ALL is uncountable — consider listing every possible language out in some table. Rows = language #; columns = Σ, \mathbb{Z}^* ; entries: include/not include

Consider diagonal language L and \bar{L} (L with answers flipped)

\gg $L \neq \emptyset$ \rightarrow $L \in ALL$

\gg $\bar{L} \neq \emptyset$ \rightarrow $\bar{L} \in ALL$

\gg $L \neq \emptyset \wedge \bar{L} \neq \emptyset \rightarrow L \in ALL$ (contradiction)

\gg Aside: similar argument to show $|P(S)| > |S|$ (power set goes \uparrow)

So, as ALL uncountable and D countable, ALL $\not\equiv$ D uncountable.

PROOF THAT L(DIA) IS UNDECIDABLE (Languages that don't accept themselves)

Let S be language of all valid Turing Machine encodings

\bullet $L(DIA) = \{ \langle M \rangle \mid M \text{ does not accept } \langle M \rangle \}$ (M does not accept itself)

Claim: For any valid TM M^* , $L(M^*) \approx L(DIA)$.

Suppose $L(DIA) = L(M^*)$. Then, $L(M^*)$ itself either $\in L$ or $\in \bar{L}$

Case 1: $M^* \in L$ ($L(DIA)$)

Then, $M^* \not\in L(M^*)$. But $L(M^*) = L(DIA)$, so contradiction!

Case 2: $M^* \in \bar{L}$ (DIA)

Then, $M^* \not\in L(M^*)$. But $L(M^*) = L(DIA)$, so contradiction!

Thus, no such M^* exists and $L(DIA)$ is thus undecidable.

LACC UNDECIDABLE (TR)

Let $LACC = \{ \langle M, x \rangle \mid M \text{ accepts } x \}$

"Universal" TM that can simulate any other.

Show $LACC \subseteq LACC$

Suppose for sake of contradiction a decider D exists for $LACC$

On input $\langle M, x \rangle$:

\gg If $D(\langle M, x \rangle) = \text{reject} \rightarrow \text{reject}$

\gg else return $M(x)$

As $LACC$ undecidable, so is $LACC$

* * REMINDERS *

ALL-PAIRS SHORTEST PATH (Floyd-Warshall) — DP

Given directed, weighted graph on n vertices Σ, \dots, Σ_n , calculate shortest path $d(i, j) \forall (i, j)$ s.t. $i \neq j$. Assume: no negative cycles

IDEA: Recursively compute $d^l(i, j) \forall (i, j)$ — shortest path using vertices that are in range $\{1, \dots, l\}$. This ensures dependent entries done

If $l = 0$, $d^0(i, j) = w(i, j)$ [could be ∞ if no edge]

Else: $d^l(i, j) = \min \{ d^{l-1}(i, k) + d^{l-1}(k, j) \}$

$d(i, j) = d^n(i, j)$

Bottom - Up — 3D Table (l, i, j)

\gg for $i, j: T[0, i, j] = w(i, j)$

\gg for $l = 1 \rightarrow n$ for $i, j: T[l-1, i, j] = \min \{ T[l-1, i, k] + T[l-1, k, j] \}$

works b/c entire table is filled already

KRUSKAL'S MINIMUM SPANNING TREE (Greedy)

G is weighted, connected, undirected graph (V, E)

Spanning Tree of G: acyclic, connected subgraph $(E' \subseteq E)$ combining all vertices

Find minimum-weight spanning tree of G

KRUSKAL'S (Greedy):

Let $T = \emptyset$ (empty graph)

Let $E = \text{sorted}(E)$ — increasing order of weight

For $e \in E$:

\gg If $T + e$ is acyclic: $T \leftarrow T + e$

return T

Proof: Kruskal finds MST

\gg Suppose T' is MST. Show that T' can be transformed to T (still optimal)

\gg Let e_1, e_2, \dots, e_{n-1} be edges of T (sorted, in order of addition)

\gg Proof by induction:

\gg Base Case: $T(e_1 \rightarrow e_n) = T'(e_1 \rightarrow e_k)$

\gg Is: Suppose $e_j \rightarrow e_m$ OK. Consider $e_{j+1} \rightarrow e_{m+1}$ in T

\gg if $(e_{j+1}, e_m) \in T'$, we are good

\gg else: adding e_{j+1} to T creates a cycle C. Then, $\exists f \in T'$ on cycle C

\gg Remove f and swap w/ e_{j+1} . This is optimal b/c as f $\not\in T$.

Weight (f) \geq Weight(e_{j+1})

O-1 Knapsack (DP)

Given Weights $W = (w_1, w_2, \dots, w_n)$ and Values $V = (v_1, v_2, \dots, v_n)$

Have knapsack with weight capacity C — pick subset of items ($\Delta \subseteq \Delta$) such that $\sum w_i \leq C$ and $\sum v_i$ is maximized

Subproblem: $K(l, i, C)$: value when only consider first i items, capacity C

\gg If $C < 0$ $\rightarrow C \rightarrow 0$

PROOF THAT 2-TAPE TM \equiv 1-TAPE

- 2-tape can simulate 1-tape by ignoring second tape
- 1-tape can simulate 2-tape as follows:
 - parity determines 1st/2nd tape
 - remember head thru special symbols 0', 1', ⊥'
 - to change state: find first head, modify tm find 2nd head, modify

DFA to decide $x \in L^*$ or $x = ab^*$

```

graph LR
    q0((q0)) -- a --> q1((q1))
    q1 -- b --> q0
    q1 -- a --> q2((q2))
    q2 -- b --> q0
    q2 -- a --> q3((q3))
    q3 -- b --> q0
    q3 -- a --> REJECT((REJECT))
  
```

TM that accepts

```

graph LR
    qstart((qstart)) -- "0 → 0, L" --> q0[0 → 0, L]
    q0 -- "1 → 1, L" --> qback((qback))
    qback -- "1 → ⊥, R" --> qstart
    qback -- "0 → 0, R" --> qscan((qscan))
    qscan -- "1 → 1, R" --> qend((qend))
    qend -- "1 → ⊥, L" --> qreject((qreject))
    qreject -- "0 → ⊥, R" --> qstart
    qstart -- "1 → 1, R" --> qreject
  
```

PROOF: L_{DIAG} is UNRECOGNIZABLE

- $L_{\text{DIAG}} = \{ \langle M \rangle \mid M \text{ is TM}, L(M) \neq \emptyset \}$
- L_{DIAG} recognizer:
 - On input $x = \langle M \rangle$ for some TM M return $M(\langle M \rangle)$;
 - If $x \in L(M)$: accept
- As L_{DIAG} undecidable and L_{DIAG} recognizable, L_{DIAG} must be unrecognizable

PROOF: L_\emptyset is UNRECOGNIZABLE

- $L_\emptyset = \{ \langle M \rangle \mid M \text{ is TM}, L(M) = \emptyset \}$
- All TMs that accept no input
- L_\emptyset is undecidable:
 - $L_{\text{DIAG}} \leq_T L_\emptyset$ (Let D decide L_\emptyset)
 - D : On input $\langle M \rangle$
 - M_1 := "return $M(\langle M \rangle)$ on all input"
 - return $D(M_1)$
- L_\emptyset is recognizable
 - On input $\langle M \rangle$:
 - simulate $M(x)$ for all x by DOWNTAILING
 - accept as soon as M accepts some x
- Thus, L_\emptyset must be unrecognizable.

PROPERTIES OF DECIDABLE LANGUAGES:

- If L is decidable, \bar{L} is decidable
- If L_1, L_2 decidable, $L_1 \cap L_2$ decidable
- If L_1, L_2 decidable, $L_1 \cup L_2$ decidable
- If L decidable, $L \leq_T L'$ for all L'
- $\forall L, L \leq_T \bar{L}$

PROPERTIES OF UNDECIDABLE LANGUAGES

- If L undecidable, \bar{L} also undecidable
- For undecidable L_1, L_2 , $(L_1 \cap L_2)$ not necessarily undecidable (consider: $L_1 \cap L_2 = \emptyset$)
- Similarly, $L_1 \cup L_2$ not necessarily undecidable (consider: $L_2 = \bar{L}_1$)
- If $A \leq_T B, B \leq_T C \rightarrow A \leq_T C$

Order inputs lexicographically

- ① Simulate one step of x_0
- ② Simulate 1 step for x_0, x_1
- ③ Simulate 1 step x_0, x_1, x_2
- ...

DOWNTAILING: Way to simulate all possible inputs on TM

RECOGNIZABILITY

- L RECOGNIZABLE if $L = L(M)$ for some M
 - M may loop on an input
- All decidable languages are recognizable
- If L_1, L_2 recognizable, so is $L_1 \cup L_2$
 - » Alternately simulate steps of M_1, M_2
 - » accept as soon as one accepts
- Similarly, $L_1 \cap L_2$ is recognizable
 - » alternately simulate recognizers — one MUST eventually accept
- COROLLARY: if L undecidable, then at least one of L, \bar{L} unrecognizable
- L is CO-RECOGNIZABLE if \bar{L} is recognizable
 - » Decidable L is both recognizable, corecognizable
 - » Countably many recognizable languages b/c countably many programs

KOLMOGOROV COMPLEXITY

- KC of a binary string x is length of shortest program that outputs x
- Complexity independent of programming language
- Implication: \exists incompressible string

PROVING UNDECIDABILITY USING Rice's

- ① Find semantic property P s.t. $L = L_P$
- ② Show P is non-trivial ($L_1 \in P, L_2 \notin P$)
- ③ Conclude L is undecidable

Ex: $L = \{ \langle M \rangle \mid M \text{ accepts all strings of length less than } 376 \}$

Let S_{376} be set of all strings less than length 376. Then let P be the set of all languages L^* s.t. $S_{376} \subseteq L^*$. Then, if $\langle M \rangle \in L$, M accepts all s.b. $s \in S_{376}$. As $S_{376} \subseteq L(M)$, $L(M) \in P$ and thus $\langle M \rangle \in L_P$. Furthermore, $Z^* \in P$ and $\emptyset \notin P$, so P is non-trivial (both recognizable). So, by Rice's, L is undecidable.

RICE'S THEOREM

- For any non-trivial property P (such that there exists recognizable $A \in P$ and $B \notin P$), $L_P = \{ \langle M \rangle \mid L(M) \in P \}$ is undecidable:
 - Only decidable when P includes all/no langs
 - P is a PROPERTY — just a set of languages
 - For any non-trivial set of languages, impossible to decide whether a given TM recognizes a language in that set
- PROOF: Show $L_{\text{HALT}} \leq_T L_P$
 - Let D decide L_P ; Construct D' for L_{HALT}
 - CASE 1: $\emptyset \in P$ (P does not include empty language)
 - D' on input (M, x)
 - Let M_1 := "On input w : simulate $M(w)$ if $M(w)$ halts: return $M_A(w)$, $A \in P$ "
 - Return $D(M_1)$
 - So, $L(M) = A \in P$ if halts, and $\emptyset \in P$ if not.
 - CASE 2: $\emptyset \notin P$
 - Some proof, substitute M_A for M_B , $!D(M)$
 - So, if M halts $\rightarrow L(M) = B \in P$
 - if M loops $\rightarrow L(M) = \emptyset \notin P$

IMPLICATIONS:

- Any non-trivial question about semantics (meaning, what it computes) is undecidable
- Behavior of a program nearly impossible

Example 4.7: Let $L = \{ \langle M \rangle \mid L(M) \text{ writes a } \perp \text{ to its tape on input } z \}$ is undecidable!

PROVE DECIDABILITY w/ Rice's

- ① Find P s.t. $L = L_P$ ($\forall M \in L, L(M) \in P$)
- ② Show P is trivial
- ③ Conclude L decidable.

Example 4.7: Let $L = \{ \langle M \rangle : L(M) \subseteq \Sigma^* \}$. Show L is decidable by applying Rice's Theorem.

Solution: Observe that $L = \{ \langle M \rangle : L(M) \subseteq \Sigma^* \}$. All languages (not just the recognizable languages) are a subset of Σ^* , so every machine should be a part of this set. What we need to do is construct a P such that L_P is the set of all machines. If we take P to be the set of all languages Σ^* , then we should have $L_P = \{ \langle M \rangle : L(M) \in \Sigma^* \}$. To show $L_P = L$

$(M \in L \iff L(M) \subseteq \Sigma^* \iff L(M) \subseteq \Sigma^* \iff \langle M \rangle \in L_P)$

Then we need to show P is trivial. Every language (not just the recognizable languages) is an element Σ^* , so all recognizable languages satisfy (or are contained in) P . Then by Rice's Theorem, L is decidable.