

NP

Satisfiability or SAT

Input: a formula ϕ over variables x_1, x_2, \dots, x_n

Solution: find a satisfying assignment or output no satisfying assignment exists.

Example: $\phi = (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_1 \vee \bar{x}_2) \wedge (x_2 \vee \bar{x}_3) \wedge (\bar{x}_2 \vee \bar{x}_3)$
try all possible assignments $\rightarrow 2^n$

VERIFY (Instance ϕ , Solution $F: \{x_1, \dots, x_n\} \rightarrow \{0, 1\}$) \rightarrow (1) $\phi = F(F(x_1)) \wedge \dots \wedge F(F(x_n))$
(2) Output $\chi_{F(\phi)}$.

P & NP

P: Class of search problems where we can find a solution in polynomial time.

HornSAT EP. & 2SAT EP

NP: Class of all search problems, i.e. the class of all problems where we can verify a solution in polynomial time. $SAT \in NP$. $2SAT \in P$. $P \subseteq NP$

Graph 3-Coloring ENP

Input: $G = (V, E)$

Solution: Coloring function $c: V \rightarrow \{R, G, B\}$, so that $c(u), c(v) \in G$ ($c(u) \neq c(v)$) or output no coloring function exists.

Proof: **VERIFY** (Instance $G = (V, E)$, Solution $c: V \rightarrow \{R, G, B\}$) \rightarrow (1) Output if $\forall u, v \in V, c(u) \neq c(v)$.
(2) else output 0.

Vertex Cover (VC) ENP

Input: $G = (V, E)$, b

Solution: $A \subseteq V$, s.t. $|A| \leq b$, $s.t. \forall (u, v) \in E$, either $u \in A$ or $v \in A$ or output no solution

Proof: **VERIFY** (Instance $G = (V, E)$, Solution $A \subseteq V$, $|A| \leq b$) \rightarrow (1) Output 0 if $\forall (u, v) \in E$, $s.t. u \notin A \wedge v \notin A$
(2) Output 1 otherwise

Factoring ENP

Input: $N = pq$, Solution p, q

Proof: **VERIFY** (Instance N , Solution p, q) \rightarrow 0 otherwise

TSP ENP

Instance: n vertices and distances d_{ij} between them & a bound B.

Solution: a permutation $\tau: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ s.t. $\sum_{i=1}^n d_{\tau(i)\tau(i+1)} \leq B$ or output no solution exists

Proof: **VERIFY** (Instance $\tau: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, Solution B) \rightarrow (1) Output 1 if $\sum_{i=1}^n d_{\tau(i)\tau(i+1)} \leq B$
(2) Output 0 otherwise

RC/HC ENP

Instance: $G = (V, E)$

Solution: permutation $\tau: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ s.t. $(\tau(i), \tau(j)) \in E$ or output no solution exists

Proof: **VERIFY** (Instance $\tau: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, Solution E) \rightarrow (1) Output 1 if $\forall (i, j) \in E, \tau(i) \neq \tau(j)$
(2) Output 0 otherwise

Composition of Reductions

Lemma: $\exists f: A \rightarrow B \wedge g: B \rightarrow C$, then $h = g \circ f$, $h_A(S) = h_B(g(f(S)))$

Reductions $A \xrightarrow{p} B$ $A \in NP$

A reduces to B in polynomial time.
An algorithm for B yields an algorithm for A.
B is at least as hard as A.

Algorithm for A

Instance I

Algo for B

Solution S for B

h(S) as the solution for I

Reduction = (f, h)
two polynomial time process

1) Running time: $F \geq h$ run in time polynomial in |I|.
2) If B outputs S as a solution to $f(I)$ (under problem B) then, $h(S)$ is a solution to I (under prob A).
3) If B outputs "no solution exists" then no solution to I exists either. (If I has a solution, then $f(I)$ also has a solution)

Rudrata Cycle \rightarrow Rudrata Half Cycle

Input: $G = (V, E)$ **Input:** $G' = (V', E')$

Solution: cycle visiting each vertex exactly once **Solution:** cycle visiting $\frac{|V|}{2}$ of the vertices exactly once

RC

RHC

C in G'

h(C) = C in G

no cycle in G'

no cycle in G

f(G): $G' = (V', E') \rightarrow E' = E \quad V' = V \cup \{n+1, \dots, 2n\}$

proof: (1) $f \circ h$ are polynomial time in |I|.
(2) If C is a RHC in G' , then $h(C)$ is also RHC in G
i) C doesn't contain $n+1, \dots, 2n$
ii) Number of vertices in C is $n = |V'|$. ($|V'| = 2|V|$)
 \Rightarrow C contains all of the vertices in $1 \dots n$
 \Rightarrow C is a RC in G .
iii) If G has a RC, then G' has a RHC.
Let C be the RC in G, then (C has |V| vertices)
C is also the RHC in G' .

SAT \rightarrow 3SAT

Input: Formula ϕ **Input:** Formula $\phi + \text{vars}$

Solution: Assignment $S: \{x_1, \dots, x_n\} \rightarrow \{0, 1\}$ **Solution:** Assignment $S: \{x_1, \dots, x_n\} \rightarrow \{0, 1\}$

$(a_1 \vee a_2 \dots \vee a_k) \rightarrow (a_1 \vee a_2 \vee a_3) \wedge (\bar{a}_1 \vee a_2 \vee a_4) \dots \wedge (\bar{a}_1 \vee \bar{a}_2 \vee a_k)$

f: $\chi_{\phi} \rightarrow \chi_{\phi'}$ **h(S):** recover a solution to ϕ

Lemma: if $w \cdot f(\phi)$ was a satisfying assignment S then $h(S)$ is a satisfying assignment for ϕ ($\exists i$ s.t. $a_i = 1$)

Lemma: if ϕ has a satisfying assignment then it also has a satisfying assignment. $\rightarrow w$ be satisfied.
Say $a_i = T$. y_1, \dots, y_{k-2} to be True and rest to be False

(S, t) Rudrata Path \rightarrow Rudrata Path

Input: $G = (V, E)$, s.t. **Input:** $G = (V, E)$

Solution: A path starting at s and ending at t that visits each vertex exactly once. **Solution:** a cycle that visits each vertex exactly once.

Independent Set \rightarrow Vertex Cover

Input: $G = (V, E)$, s.t. **Input:** $G = (V, E)$, b

Solution: $S \subseteq V$, s.t. $|S| = g$ **Solution:** $S \subseteq V$, s.t. $|S| = b$

$\exists \forall (u, v) \in E, \text{ either } u \in S \text{ or } v \in S$ $\wedge \forall (u, v) \in E, \text{ either } u \in S \text{ or } v \in S$

f(G, S, t) = G'

$G' = (V', E')$ $V' = V \setminus S$

$E' = E \setminus \{(x, y) | (x, y) \in E \wedge x \in S \wedge y \in S\}$

$h(S) = C = \{x \in V | \forall y \in S, (x, y) \in E\}$

Proof: (1) Runtime of f and h
(2) If S is a RC in G' then $h(S)$ is a RP-(s, t) in G .
(3) If G has a RC - RP, then G' has a RC.

Circuit SAT ENP

Instance: a boolean circuit C A DAG with 5 kinds of gates.

Solution: an assignment to unknown input gates s.t.
1) AND & OR gates of indegree 2.
2) NOT gate of indegree 1.
output gate outputs TRUE.
3) known input gates
4) unknown input gates

3D Matching

Input: n boys, girls and pets with preference tuples $\{(b, g, p)\}$

Solution: n disjoint tuples.

3SAT \rightarrow 3D Matching

Instance: Formula ϕ on variable x_1, \dots, x_n **Input:** n boys, girls and pets with preference tuples $\{(b, g, p)\}$

Solution: A satisfying assignment $S: \{x_1, \dots, x_n\} \rightarrow \{0, 1\}$ **Solution:** n disjoint tuples.

Gadgets

gates

branch

Input: b_1, p_1, b_2, p_2 \rightarrow (b_1, p_1, b_2)
 p_1, g_1, p_2, b_2 \rightarrow (p_1, g_1, p_2)
 $(b_1, g_1, p_1) \wedge (b_2, g_2, p_2) \rightarrow (b_1 \wedge b_2, g_1 \wedge g_2, p_1 \wedge p_2)$

bc = $(x \vee y \vee z)$, create bc, gc.
 $\Rightarrow x = \text{true}$ or $y = \text{false}$ or $z = \text{true}$.
 $(bc, jc, pc, pc) \rightarrow (bc, pc, pc)$
 $\text{or } (bc, jc, pc) \text{ or } (bc, jc, pc) \rightarrow (bc \wedge jc, pc, pc)$.

Zero-One Equations

Instance: $A \in \{0, 1\}^{m \times n}$ **Input:** A

Solution: $x \in \{0, 1\}^n$ s.t. $Ax = 1$ **Input:** A

3D Matching \rightarrow ZOE

Input: t - boys, girls, pets **Input:** $A \in \{0, 1\}^{m \times n}$ with preference tuples $\{(b, g, p)\}$. **Solution:** $x \in \{0, 1\}^n$ s.t. $Ax = 1$

Solution: t disjoint tuples

all tuple where 1st boy is included. $x_i = 1$ if i th T is part of the 3DM solution.

part of the 3DM solution.

ZOE \rightarrow RC WPE

Instance: $A \in \{0, 1\}^{m \times n}$ **Input:** A

Solution: $x \in \{0, 1\}^n$ s.t. $Ax = 1$ **Input:** A

RC WPE \rightarrow RC

Instance: $G = (V, E)$ **Input:** $G = (V, E)$

Solution: $RC \subseteq V$ s.t. $\forall (u, v) \in E, \text{ either } u \in RC \text{ or } v \in RC$ **Input:** $G = (V, E)$

Solution: Cycle S Visiting each vertex exactly one.

RC \rightarrow TSP

Instance: $G = (V, E)$ Instance: distances d_{ij} & a tour D .
 Solution: Cycle S visiting each vertex exactly once
 with exactly one visit to v_i .

$d_{ij} = 1 \text{ if } (i,j) \in E$
 $d_{ij} = \infty \text{ if } (i,j) \notin E$

ZOE \rightarrow Subset Sum

Instance: $A = \{a_1, a_2, \dots, a_n\}$ Instance: a_1, a_2, \dots, a_n, w
 Solution: $x_1, x_2, \dots, x_n \in \{0, 1\}$ s.t.
 $\sum_i a_i x_i = w$

$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \rightarrow a_i = \sum_j a_{ij} x_j$
 $w = \sum_i a_i x_i$

Up to Now:
 All of NP
 SAT
 3SAT

Independent Set
 Vertex cover
 Clique
 RC
 ZOE
 Subset Sum

NP-Complete problem still need a solution
 ① Intelligent "exponential search" → running time could be exponential
 ② Approximation Algorithms → polynomial → practical instances
 ③ Heuristic → no guarantees on the runtime or optimality of the solution.

Backtracking

Start with some problem P_0
 Let $S = \{P_0\}$, the set of active subproblems
 Repeat while S is not empty:
 choose subproblem P_i , and remove it from S .
 expand into smaller subproblems P_1, \dots, P_k .
 For each P_i
 if $\text{test}(P_i)$ succeeds: halt and announce this solution
 test(P_i) fails: discard P_i
 otherwise: add P_i to S .
 Announce that there is no solution.

Branch and Bound: Generalizing backtracking to search problems

Start with some problem P_0
 Let $S = \{P_0\}$, the set of active subproblems
 bestsofar = ∞
 Repeat while S is not empty:
 choose subproblem P_i (partial solution) $\in S$ and remove it from S .
 expand into smaller subproblems P_1, \dots, P_k .
 For each P_i
 if P_i is a complete solution: update bestsofar
 else if lowerbound(P_i) < bestsofar: add P_i to S .
 Return bestsofar.

TSP - branch and bound

$W_{TSP} = d_{12} + d_{23} + \dots + d_{n1}$ is minimized
 Lemma: $W_{TSP} \geq W_{MST}$

Approximation Algorithms for an optimization problem

Instance $I \rightarrow OPT(I)$ approx ratio $\alpha = \max_I \frac{OPT(I)}{OPT(I)}$ (minimized)
 (minimization problem e.g. TSP) $\rightarrow A$ $A = \max_I \frac{OPT(I)}{OPT(I)}$ (maximized)

Set Cover

Input: A set of elements B ; sets $S_1, S_2, \dots, S_m \subseteq B$
 Output: Smallest selection of S_i , whose union is B

Greedy Algorithm: Repeat until all elements of B are covered
 Pick the set S_i with the largest number of uncovered elements.

Claim: say $|B|=n$ and $OPT(I)=k$. Then greedy algorithm uses at most $k \ln n$ sets

proof. let n_t be the number of uncovered elements in B after t iterations of our greedy algorithm.
 at most one of these sets has $\geq \frac{n_t}{k}$ uncovered elements.
 $\therefore n_{t+1} \leq n_t - \frac{n_t}{k} = n_t(1 - \frac{1}{k}) \Rightarrow n_t \leq n(1 - \frac{1}{k})^t \Rightarrow n_t \leq n e^{-t/k}$
 $\therefore t \geq k \ln n$

Local Search Heuristics

Let S be any initial solution
 while there is some solution S' in the neighborhood of S for which $cost(S') < cost(S)$: replace S with S'
 returns S

Vertex Cover

Input: $G = (V, E)$ Output: Subset $S \subseteq V$ such that $|S|$ is minimized and S touches every edge.

approx solution: ① find a maximal matching $M \subseteq E$
 ② Return $S = \{ \text{all endpoint of edges in } M \}$.
 $|S| = 2|M| \leq 2 \times \text{size of any VC}$

Clustering

Input: Points $X = \{x_1, \dots, x_n\}$, with distance metric $d(\cdot, \cdot)$, integer k .
 Output: k clusters C_1, \dots, C_k s.t. max max $d(x_i, C_j)$
 minimize the diameter of clusters $\sum_j \text{diam}(C_j)$
 $\text{diameter of } C_j = \max_{x_i, x_j \in C_j} d(x_i, x_j)$

Pick any input $x_i \in X$ as the first cluster center
 for $i=2 \dots k$
 let x_i be the point farthest from x_1, \dots, x_{i-1}
 create k clusters: $C_i = \{x_j \mid x_j \text{ closest to } x_i\}$

Lemmas: $b_i \cdot \text{diam}(C_i) \leq d(x_i, x_j) + \text{diam}(C_i)$
 proof: $d_{ij} = d(x_i, x_j) \leq d(x_i, x_i) + d(x_i, x_j) = r_i + \text{diam}(C_i)$ (diameter)

RC \rightarrow TSP

Input: $G = (V, E)$ Instance: distances d_{ij}
 Solution: cycles visiting V : $S_1 \rightarrow S_2 \rightarrow \dots \rightarrow S_m \rightarrow S_1$ s.t. each vertex exactly once
 $d_{12} + d_{23} + \dots + d_{m1} = \text{cost}$

$d_{ij} = 1 \text{ if } (i,j) \in E$
 $d_{ij} = \infty \text{ if } (i,j) \notin E$

if G has ARL $\Rightarrow G'$ has a TSP solution of cost n
 if G does not have ARL $\Rightarrow G'$ has no TSP solution of cost n
 $ARL \rightarrow 2\text{-TSP}$
 when $n=d_m$. if G has ARL $\Rightarrow G'$ has a TSP solution of cost n
 2. if G doesn't have ARL \Rightarrow at least $nd_m = n(n+2)$

2-TSP

Instance: distances d_{ij}
 Solution: a permutation $T: S_1 \rightarrow S_2 \rightarrow \dots \rightarrow S_m \rightarrow S_1$ s.t.
 $d_{12} + d_{23} + \dots + d_{m1} = \text{cost}$

$d_{ij} = d_{ji}$ $d_{ij} + d_{jk} \geq d_{ik}$

Lemma: ① MST can be a good starting point
 ② $d_{MST} \leq d_{2\text{-TSP}}$
 ③ $d_{2\text{-TSP}} \leq d_{MST}$

Knapsack (without rep) **Dynamic**

Input: n, w_i, v_i Goal: Most valuable combination with total weight $\leq w$
 we will give an approx algorithm s.t. $k = O(\log n)$

Discard any items with weights $> w$
 let $U_{max} = \max_i w_i$
 Rescale $\hat{w}_i = \lfloor w_i / U_{max} \rfloor$
 Run DP algorithm with values $\{\hat{w}_i\}$.
 Output the resulting choice of items

Set $\sum_i \hat{v}_i = \sum_i \lfloor v_i / U_{max} \rfloor \geq \sum_i (\frac{v_i}{U_{max}}) \geq \frac{\sum_i v_i}{U_{max}} = 1$

GI

$G_0 \approx G_1$, $G_0 \rightarrow V$

P $\xrightarrow{\pi} V$

zero-knowledge: If $G_0 \approx G_1$, then V learns nothing more than the fact that $G_0 \approx G_1$.
 $\hookrightarrow V$ will be able to generate the interaction on his own.

$G_0 \xrightarrow{\pi} P$
 $\circledcirc H = \neg(G_0) \rightarrow V$
 $b=1 \xrightarrow{\phi \rightarrow V} \circledcirc b$
 $b=0 \xrightarrow{\phi = \pi} \circledcirc H$
 $\hookrightarrow \text{if } \phi \wedge (G_0) = H \text{ then output } 1 \text{ else } 0$

Soundness: $G_0 \neq G_1$, $H \xrightarrow{\phi} G_0$, $H \approx G_1$
 with probability $\frac{1}{2}$ $b+H$ then P will have ∞ way to make V output 1.

zero-knowledge: $G_0 \approx G_1$, $\circledcirc H = \neg(G_0) \rightarrow V$
 $\circledcirc H = \neg(G_0) \approx \neg(G_1)$

Simulated Annealing (introduce temperature parameter T).
 Let S be any initial solution
 Randomly choose a solution S' in the neighborhood of S
 if $cost(S') < cost(S)$: replace S by S'
 else
 : replace S by S' with probability
 $e^{-\frac{|cost(S') - cost(S)|}{T}}$

Thinking of NP as a proof (Interactive Proof)

power $P \xrightarrow{\phi} V$ verify. \rightarrow polytime.
 completeness: $H = \phi$ is true. then in $P(\phi, w) \rightarrow V(H)$ outputs 1
 soundness: $\neg H = \phi$ is true. then in $P(\neg\phi, w) \rightarrow V(H)$ outputs 1 with very small probability e.g. $2^{-1000000}$ parameters

Verify $A \vee B = C$

$\bullet 1 \leftarrow 1, \dots, q \quad \vec{r} = (1, 1, \dots, r^{n-1})$
 $\bullet C \times \vec{r} = (A \vee B) \wedge \vec{r} \quad O(n^3)$ faster than $O(n^6)$
 O(n³)

Soundness: $D = A \vee B$ CTD

then $\exists i: c_i \neq d_i \wedge c_i \cdot \vec{r} = d_i \cdot \vec{r} \Rightarrow C(c_i - d_i) \cdot \vec{r} = 0$
 $(P_1, P_2, \dots, P_m)(1, r, \dots, r^n) = 0$. $O(\frac{1}{r^3})$, $m \gg n$
 $\therefore P(\vec{r}) = P_1 \wedge P_2 \wedge \dots \wedge P_m \vec{r}^m = 0 \Rightarrow P_1, P_2, \dots, P_m \vec{r}^m = 0$

Graph Isomorphism

Instance $G_0 = (V_0, E_0)$ $G_1 = (V_1, E_1)$
 Solution: $\pi: V_0 \rightarrow V_1$ s.t. $b \in \text{unis}(E_0)$ iff $(\pi(u), \pi(v)) \in E_1$
 $H \xrightarrow{G_0, P} \circledcirc H = \neg(G_0) \rightarrow V$
 $\circledcirc H = \neg(G_0) \rightarrow V$
 $b=1 \xrightarrow{\phi \rightarrow V} \circledcirc b$
 $b=0 \xrightarrow{\phi = \pi} \circledcirc H$
 completeness: $G_0 \approx G_1 \rightarrow H \approx G_1$
 soundness: $G_0 \neq G_1 \rightarrow H \approx G_1$
 $\hookrightarrow b=1 \text{ then output } 1 \quad P \xrightarrow{H = \neg(G_0)} \circledcirc b$
 $\hookrightarrow b=0 \text{ then output } 0 \quad P \xrightarrow{H = \neg(G_0)} \circledcirc b$



