

lec04: Interactive Proofs (IP)

- 1 Interactive Proofs
- 2 IP for Graph Non-Isomorphism
- 3 IP for #SAT.

1 Interactive Proofs

Q: What is a proof for?

A: Someone convinces someone else that something is true

"prover"
"verifier"
"statement"

By convention, we write all statements as set membership

"instance"

"language"

~~x is a prime~~

$x \in \mathcal{L}_{\text{primes}}$

~~15 is biprime~~

$15 \in \{pq : p, q \text{ are prime}\}$

~~ϕ is a satisfiable formula~~

$\phi \in \{\text{formula } \phi \mid \exists x \phi(x) = 1\}$

~~ϕ has 8 satisfying assignments~~

$(\phi, 8) \in \{\dots\}$

Not all statements are equally hard to prove

to prove 15 is biprime,

give 3&5

to prove 83 is prime...

?

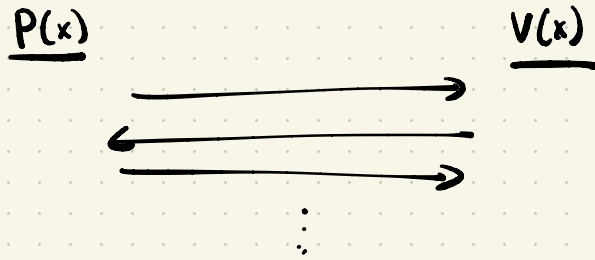
the # of satisfying assignments.. ???

Today, we define interactive proofs (IPs):

An IP is two interactive algorithms (P, V)

- V is PPT
- P is randomized, but need not be polytime
- both take input x (the instance)
- V outputs 1 (accept or reject)

- They interact like this:



- the message count & sizes must be $\text{poly}(\lambda)$, since V is.
- $\langle P, V \rangle(x)$ denotes V 's output at the end

An IP must be complete + sound for a target language $\mathcal{L}$:

Complete: $\forall x \in \mathcal{L} \quad \Pr_{P,V}[\langle P, V \rangle(x) = 1] \geq 1 - \text{completeness error}$

Sound: $\forall x \notin \mathcal{L} \quad \Pr_{P,V}[\langle P, V \rangle(x) = 1] \leq \text{soundness error}$

A special case: non-interactive proofs (NP), in which P sends one message

NP is well-studied and there are NP proofs for many hard problems:

- Hamiltonian cycle
- Circuit satisfiability (SAT)
- Graph isomorphism

But not all problems:

- Graph non-isomorphism (GNI) ✓
- #SAT: counting circuit assignments ✓
- Optimal chess (PSPACE) ✓
- Halting problem

IPs:

1. solve many of these •
2. often have better concrete efficiency

Today:

- IP for GNI
- IP for #SAT
- See a complexity course or the readings for $IP = PSPACE$
- See the "succinctness" unit for concrete efficiency.

[2] IP for GNI (Graph Non-Isomorphism)

Notation: Let $G \subseteq [n] \times [n]$ denote a directed graph on n nodes. $(a, b) \in G$ if $a \mapsto b$ in G .
 $\{1, \dots, n\}$

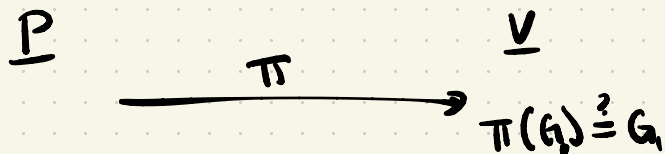
Definition G_0 and G_1 are isomorphic ($G_0 \cong G_1$) if there exists $\pi \in \text{Perms}[n]$ s.t. $\pi(G_0) = G_1$, where

all bijections
on $[n]$

$$\pi(\{(a_i, b_i)\}_i) \triangleq \{(\pi(a_i), \pi(b_i))\}_i$$

π "relabels" G 's vertices

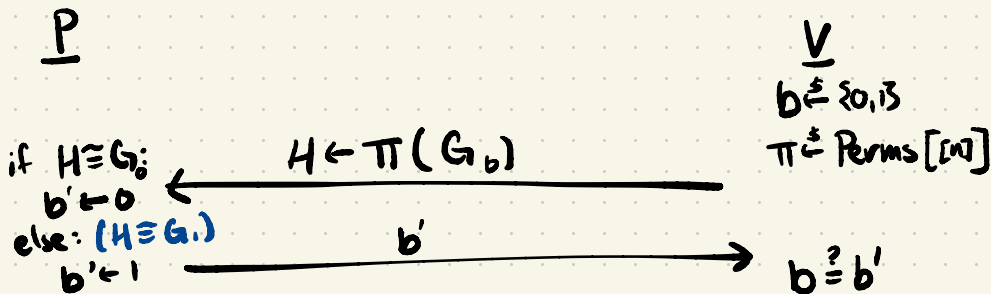
Proving $G_0 \cong G_1$ is easy. It's in NP.



What about proving $G_0 \not\cong G_1$? (GNI)

- This is not known to be in NP.
- We will give a simple IP.
 - The idea is for V to:
 1. hold the G_i behind its back
 2. choose one at random
 3. randomly scramble it
 4. show the scrambling to P
 5. have P "guess" which G_i was scrambled.

Protocol:



This IP is perfectly complete.

Pf: Fix $G_0 \not\cong G_1$,

Case $b=0$. Then, $H \cong G_0$. Then, $b'=0$. Then, V accepts.

Case $b=1$. Then, $H \cong G_1$. Then, $H \not\cong G_0$. Then, $b'=1$.

Then, V accepts.

This IP is $1/2$ -sound.

That is, for $x \notin \mathcal{L}$, $\Pr[\langle P, V \rangle(x) = 1] \leq 1/2$.

Pf. Fix $G_0 \cong G_1$, "is independent of"

Then, $H \perp\!\!\!\perp b$. Intuitively, b/c H is distributed the same, either way.

Why? Since $\exists \sigma, \sigma(G_0) = G_1$, for any H'

$$\Pr[H = H' \mid b = 1] = \Pr[\pi(\sigma(G_0)) = H']$$

$$= \Pr_{\pi}[(\pi\sigma)(G_0) = H']$$

$\uparrow$ uniform $\nwarrow$ fixed

$$= \Pr_{\pi'}[\pi'(G_0) = H']$$

$\uparrow$ uniform

$$= \Pr[H = H' \mid b = 0]$$

Then, $b' \perp\!\!\!\perp b$ (P computes b' from H)

Then, $\Pr[b' = b] = 1/2$.

• Recall that our definition of soundness required error $\leq 1/3$. So, repeat this IP twice!

$$(1/2)^2 = 1/4 \leq 1/3!$$

3 #SAT in IP

SAT : does a Boolean formula ϕ has a solution

#SAT : how many solutions?

$(\phi, K) \in \#SAT$ if ϕ has K solutions

We'll see an IP for #SAT, called Sumcheck.

Sumcheck is very important. Today we'll see its beautiful theory. Later in the course we'll see its practical applications.

Let ϕ be in 3-conjunctive normal form: (3CNF):

$$\phi = (x_1 \vee x_3 \vee \neg x_5) \\ \wedge (x_2 \vee x_3 \vee x_5) \\ \wedge \dots$$

} clauses

↑ literals:
 x_i or $\neg x_i$

m variables

n clauses : $\phi = c_1 \wedge \dots \wedge c_n$

fix a prime $p > 2^m$, $\log p = \text{poly}(m)$

we'll map a formula ϕ over $\vec{x} = x_1, \dots, x_m$ into
a polynomial $[\![\phi]\!]$ over $\vec{X} = X_1, \dots, X_m$ in $\mathbb{Z}_p$

$$[\![\text{true}]\!] = 1, \quad [\![\text{false}]\!] = 0$$

$$[\![x_i]\!] = X_i$$

$$[\![\neg \phi]\!] = 1 - [\![\phi]\!]$$

$$[\![\phi \wedge \psi]\!] = [\![\phi]\!] \cdot [\![\psi]\!]$$

$$[\![\phi \vee \psi]\!] = [\![\neg(\neg\phi \wedge \neg\psi)]\!] = 1 - (1 - [\![\phi]\!])(1 - [\![\psi]\!])$$

Ex: $[\![x_1 \vee x_3 \vee \neg x_5]\!] = 1 - (1 - X_1)(1 - X_3)X_5$

$$[\![c_1 \wedge \dots \wedge c_n]\!] = \prod_i [\![c_i]\!]$$

Key fact : $\phi(\vec{x})$ is true $\Leftrightarrow [\![\phi]\!](\vec{x}) = 1$

Now : ϕ has k solutions

$\Leftrightarrow$

$$K = \sum_{\substack{b_1 \in \mathbb{B} \\ \mathbb{B} = \{0,1\}}} \sum_{b_2 \in \mathbb{B}} \dots \sum_{b_m \in \mathbb{B}} [\![\phi]\!](b_1, b_2, \dots, b_m)$$

$\deg [\![\phi]\!] \leq 3n$

We'll forget Φ , and study this a polynomial problem
 $f \in \mathbb{Z}_p^{\leq d}[X_1, \dots, X_m]$ $m, d, \log p = \text{poly}(\lambda)$

Protocol to show $K = \sum_{b_1 \in \mathbb{B}} \dots \sum_{b_m \in \mathbb{B}} f(b_1, \dots, b_m)$:
Sumcheck Protocol

if $m=0$, then V just checks it... $K \stackrel{?}{=} f$
 if $m \geq 1$, P

Note: The claim only evaluates f on $\mathbb{B}$, but the protocol will use all of $\mathbb{Z}_p$!

computes

$$h(X_1) = \sum_{b_2 \in \mathbb{B}} \dots \sum_{b_m \in \mathbb{B}} f(X_1, b_2, \dots, b_m)$$

$$\deg h \leq d$$

$$\begin{array}{ccc} & h(X) & \\ \xrightarrow{\hspace{1cm}} & & h(0) + h(1) \stackrel{?}{=} K \\ & r & \\ \xleftarrow{\hspace{1cm}} & & r \in \mathbb{Z}_p \end{array}$$

recurse on

$$h(r) = \sum_{b_2 \in \mathbb{B}} \dots \sum_{b_m \in \mathbb{B}} f(r, b_2, \dots, b_m)$$

$$K' = h(r)$$

$$f'(X_2, \dots, X_m)$$

$$= f(r, X_2, \dots, X_m)$$

(i.e., plug $X_1 \mapsto r$ in)

Efficiency for V :

in m rounds: evaluates h (univariate) thrice.

at end: evaluates f (multivariate) once

much better than 2^m evaluations

This is perfectly complete: (induct on m)

base case: V checks directly

recursive:

$$h(x_1) = \sum_{b_2} \dots \sum_{b_m} f(x_1, b_2, \dots, b_m)$$

so:

$$h(0) + h(1) = \sum_{b_1} \dots \sum_{b_m} f(b_1, b_2, \dots, b_m) \\ = K$$

That shows V 's test passes. ✓

We also have to show the recursive claim holds

$$h(r) = \sum_{b_2} \dots \sum_{b_m} f(r, b_2, \dots, b_m)$$

this is true by h 's definition. ✓

Soundness error is $1 - (1 - d/p)^m \approx 1 - (1 - dm/p) = dm/p$ poly(λ)

If claim is false, V rejects w/ prob. $\geq (1 - d/p)^m$ exp(λ)

Induct on m :

negl(λ)

Base case: V checks directly, and rejects ✓

Recursive: Say $K \neq \sum_{b_1} \dots \sum_{b_m} f(b_1, \dots, b_m)$

Let P send $\tilde{h}(x)$.

If $\tilde{h}(x) = h(x_1)$, then $\tilde{h}(x_1) + \tilde{h}(x_2) \neq K$.

$\Rightarrow V$ rejects ✓

If $\tilde{h}(x_1) \neq h(x_1)$, then

• Fact: the recursive claim is false

w/ probability $\geq (1 - d/p)$.

• This fact finishes the proof:

$$\begin{aligned} & \Pr[V \text{ rejects}] \\ & \geq \Pr[\text{rec. claim false}] \cdot \Pr[V \text{ rec. rejects}] \\ & = (1 - d/p) \cdot (1 - d/p)^{m-1} \quad \leftarrow \text{(induction!)} \\ & = (1 - d/p)^m \end{aligned}$$

• So, let's prove the fact:

$$\begin{aligned} & \Pr[\text{rec. claim false}] \\ & = \Pr[\tilde{h}(r) \neq \sum_{b_2} \dots \sum_{b_m} f(r, b_2, \dots, b_m)] \\ & = \Pr[\tilde{h}(r) \neq h(r)] \\ & = \Pr[(\tilde{h} - h)(r) \neq 0] \\ & \geq 1 - d/p \quad \begin{array}{l} (\tilde{h} - h \text{ has } \leq d \text{ roots!}) \\ \text{(fundamental theorem} \\ \text{of algebra)} \end{array} \quad \blacksquare \end{aligned}$$

What does sumcheek look like, unrolled?

P

V

$$h_1(x_1) = \sum_{b_2} \dots \sum_{b_m} f(x_1, b_2, \dots, b_m)$$

$$\xrightarrow{h_1} h_1(0) + h_1(1) \stackrel{?}{=} k$$

$$\leftarrow r_1 \in \mathbb{Z}_p$$

$$(\text{claim: } h_1(r_1) = \sum_{b_2} \cdots \sum_{b_m} f(r_1, b_2, \dots, b_m))$$

$$h_2(x_2) = \sum_{b_3} \cdots \sum_{b_m} f(r_1, x_2, b_3, \dots, b_m)$$

$$\xrightarrow{h_2} \quad h_2(0) + h_2(1) = h_1(r_1)$$

$$\leftarrow r_2 \in \mathbb{Z}_p$$

$$(\text{claim: } h_2(r_2) = \sum_{b_3} \cdots \sum_{b_m} f(r_1, r_2, b_3, \dots, b_m))$$

$$h_3(x_3) = \sum_{b_4} \cdots \sum_{b_m} f(r_1, r_2, x_3, b_4, \dots, b_m)$$

$$\xrightarrow{h_3}$$

the last claim

$$\vdots$$

$$\rightarrow h_m(r_m) \stackrel{?}{=}$$

$$\rightarrow f(r_1, \dots, r_m)$$

single evaluation of f .