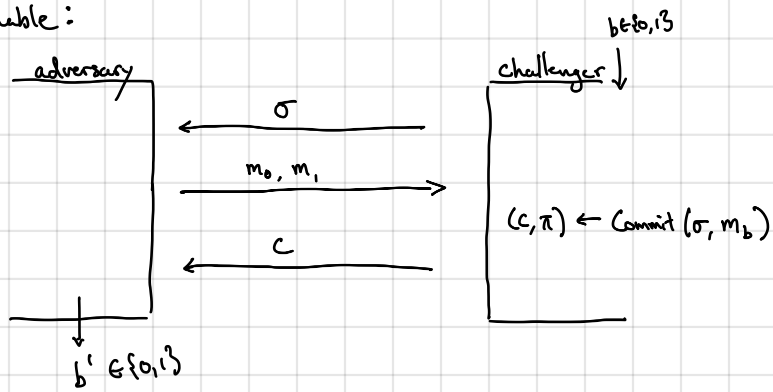
Requirements:

- Correctness: for all messages m:

$$\Pr[\sigma \leftarrow \text{Setup},\ (c,\pi) \leftarrow \text{Commit}(\sigma, m);\ \text{Verify}(\sigma, c, m, \pi) = 1] = 1$$

- Hiding: for all common reference strings $\sigma \in \{0,1\}^n$ and all efficient A, following distributions are computationally indistinguishable:

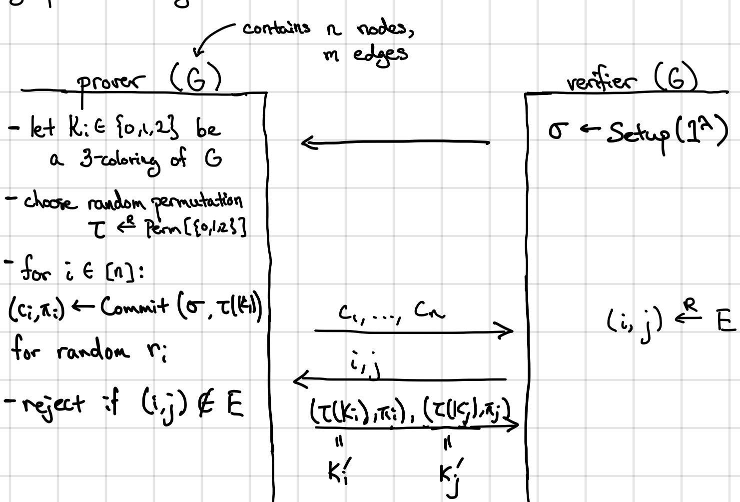


$$\left|\Pr[b' = 1 \mid b = 0] - \Pr[b' = 1 \mid b = 1]\right| = \text{negl}(\lambda)$$

- Binding: for all adversaries A, if $\sigma \leftarrow \text{Setup}$, then

$$\Pr[(m_0, m_1, c, \pi_0, \pi_1) \leftarrow A : m_0 \neq m_1 \text{ and } \text{Verify}(\sigma, c, m_0, \pi_0) = 1 = \text{Verify}(\sigma, c, m_1, \pi_1)] = \text{negl}$$

A ZK protocol for graph 3-coloring:



accept if $K_i' \neq K_j'$ and $K_i', K_j' \in \{0,1,2\}$

$\text{Verify}(\sigma, c_i, K_i', \pi_i) = 1 = \text{Verify}(\sigma, c_j, K_j', \pi_j)$

reject otherwise

Intuitively: Prover commits to a coloring of the graph
Verifier challenges prover to reveal coloring of a single edge
Prover reveals the coloring on the chosen edge and opens the entries in the commitment

Completeness: By inspection [if coloring is valid, prover can always answer the challenge correctly]

except with prob. $1-\text{negl}$.

Soundness: Suppose G is not 3-colorable. Let $K_1, \dots, K_n$ be the coloring the prover committed to. If the commitment scheme is statistically binding, $c_1, \dots, c_n$ uniquely determine $K_1, \dots, K_n$. Since G is not 3-colorable, there is an edge $(i,j) \in E$ where $K_i = K_j$ or $i \notin \{0,1,2\}$ or $j \notin \{0,1,2\}$. [Otherwise, G is 3-colorable with coloring $K_1, \dots, K_n$.] Since the verifier chooses an edge to check at random, the verifier will choose (i,j) with probability $1/|E|$. Thus, if G is not 3-colorable,

$$\Pr[\text{verifier rejects}] \geq \frac{1}{|E|}$$

Thus, this protocol provides soundness $1 - \frac{1}{|E|}$. We can repeat this protocol $O(|E|^2)$ times sequentially to reduce soundness error to

$$\Pr[\text{verifier accepts proof of false statement}] \leq \left(1 - \frac{1}{|E|}\right)^{|E|^2} \leq e^{-|E|} = e^{-m} \quad [\text{since } 1 + x \leq e^x]$$

Zero Knowledge: We need to construct a simulator that outputs a valid transcript given only the graph G as input.
Let V^* be a (possibly malicious) verifier. Construct simulator S as follows:

1. Run V^* to get σ^*.
2. Choose $K_i \leftarrow \{0,1,2\}$ for all $i \in [n]$.
   Let $(c_i, \pi_i) \leftarrow \text{Commit}(\sigma^*, K_i)$
   Give $(c_1, \dots, c_n)$ to V^*.

   $\Big\}$ simulator does not know coloring so it commits to a random one
3. V^* outputs an edge $(i,j) \in E$
4. If $K_i \neq K_j$, then S outputs (K_i, K_j, π_i, π_j).
   Otherwise, restart and try again (if fails λ times, then abort)

Simulator succeeds with probability $2/3$ (over choice of $K_1, \dots, K_n$). Thus, simulator produces a valid transcript with prob. $1 - \frac{1}{3^\lambda} = 1 - \text{negl}(\lambda)$ after λ attempts. It suffices to show that simulated transcript is indistinguishable from a real transcript.

- Real scheme: prover opens K_i, K_j where $K_i, K_j \xleftarrow{R} \{0,1,2\}$ [since prover randomly permutes the colors]
- Simulation: K_i and K_j sampled uniformly from $\{0,1,2\}$ and conditioned on $K_i \neq K_j$, distributions are identical

In addition, (i,j) output by V^* in the simulation is distributed correctly since commitment scheme is computationally-hiding (e.g. V^* behaves essentially the same given commitments to a random coloring as it does given commitment to a valid coloring

If we repeat this protocol (for soundness amplification), simulator simulate one transcript at a time

Summary: Every language in NP has a zero-knowledge proof (assuming existence of PRGs)

PRGs imply commitments

In many cases, we want a stronger property: the prover actually "knows" <u>why</u> a statement is true (e.g., it knows a "witness")

For instance, consider the following language:

$$\mathcal{L} = \{h \in \mathbb{G} \mid \exists x \in \mathbb{Z}_p : h = g^x\} = \mathbb{G}$$

$\mathbb{G}$: group of order p; g: generator of $\mathbb{G}$

Note: this definition of $\mathcal{L}$ implicitly defines an NP relation $\mathcal{R}$:
$\mathcal{R}(h, x) = 1 \iff h = g^x \in \mathbb{G}$

In this case, all statements in $\mathbb{G}$ are true (i.e., contained in $\mathcal{L}$), but we can still consider a notion of proving <u>knowledge</u> of the discrete log of an element $h \in \mathbb{G}$ — conceptually <u>stronger</u> property than proof of membership

<u>Philosophical question</u>: What does it mean to "know" something?

If a prover is able to convince an honest verifier that it "knows" something, then it should be possible to <u>extract</u> that quantity from the prover.

<u>Definition.</u> An interactive proof system (P, V) is a proof of knowledge for an <u>NP relation $\mathcal{R}$</u> if there exists an efficient extractor $\mathcal{E}$ such that for any x and any prover P^*

proof of knowledge is parameterized by a specific <u>relation</u> $\mathcal{R}$ (as opposed to the language $\mathcal{L}$)

$$\Pr[w \leftarrow \mathcal{E}^{P^*}(x) : \mathcal{R}(x, w) = 1] \geq \Pr[\langle P^*, V \rangle (x) = 1] - \varepsilon$$

more generally, could be polynomially smaller

ε: knowledge error

Trivial proof of knowledge: prover sends witness in the <u>clear</u> to the verifier
↳ In most applications, we <u>additionally</u> require zero-knowledge

<u>Note</u>: knowledge is a <u>strictly</u> stronger property than soundness
↳ if protocol has knowledge error $\varepsilon \Rightarrow$ it also has soundness error ε (i.e. a dishonest prover convinces an honest verifier of a false statement with probability at <u>most</u> ε)

<u>Proving knowledge of discrete log (Schnorr's protocol)</u>

Suppose prover wants to prove it knows x such that $h = g^x$ (i.e. prover demonstrates knowledge of discrete log of h base g)

assume $g, h \in \mathbb{G}$ where $\mathbb{G}$ has prime order q

Prover		Verifier
$r \xleftarrow{R} \mathbb{Z}_p$		
$u \leftarrow g^r$	$\xrightarrow{\quad u \quad}$	
	$\xleftarrow{\quad c \quad}$	$c \xleftarrow{R} \mathbb{Z}_p$
$z \leftarrow r + cx$	$\xrightarrow{\quad z \quad}$	

verify that $g^z = u \cdot h^c$

Completeness: if $z = r + cx$, then

$$g^z = g^{r+cx} = g^r g^{cx} = u \cdot h^c$$

zero knowledge only required to hold against an honest verifier (e.g., view of the honest verifier can be simulated)

Honest-Verifier Zero-Knowledge: build a simulator as follows (familiar strategy: run the protocol in "reverse"):

on input (g, h):

1. Sample $z \xleftarrow{R} \mathbb{Z}_p$
2. Sample $c \xleftarrow{R} \mathbb{Z}_p$
3. set $u = g^z / h^c$ and output (u, c, z)

uniformly random challenge

uniformly random group element since z is uniformly random

chosen so that $g^z = u \cdot h^c$ (relation satisfied by a valid proof)

Simulated transcript is identically distributed as the real transcript with an honest verifier

What goes wrong if the challenge is not sampled uniformly at random (i.e., if the verifier is dishonest)

Above simulation no longer works (since we cannot sample z first)

↳ To get general zero knowledge, we require that the verifier first commit to its challenge (using a statistically hiding commitment)

Knowledge: Suppose P^* is (possibly malicious) prover that convinces honest verifier with probability 1. We construct an extractor as follows:

for simplicity, we assume P^* succeeds with probability 1

1. Run the prover P^* to obtain an initial message u.
2. Send a challenge $c_1 \xleftarrow{R} \mathbb{Z}_p$ to P^*. The prover replies with a response z_1.
3. "Rewind" the prover P^* so its internal state is the same as it was at the end of Step 1. Then, send another challenge $c_2 \xleftarrow{R} \mathbb{Z}_p$ to P^*. Let z_2 be the response of P^*.
4. Compute and output $x = (z_1 - z_2)(c_1 - c_2)^{-1} \in \mathbb{Z}_p$.

Since P^* succeeds with probability 1 and the extractor perfectly simulates the honest verifier's behavior, with probability 1, both (u, c_1, z_1) and (u, c_2, z_2) are both accepting transcripts. This means that

$$g^{z_1} = u \cdot h^{c_1} \quad \text{and} \quad g^{z_2} = u \cdot h^{c_2}$$

$$\Rightarrow \frac{g^{z_1}}{h^{c_1}} = \frac{g^{z_2}}{h^{c_2}} \Rightarrow g^{z_1 + c_2 x} = g^{z_2 + c_1 x}$$

$$\Rightarrow x = (z_1 - z_2)(c_1 - c_2)^{-1} \in \mathbb{Z}_p$$

with overwhelming probability, $c_1 \neq c_2$

Thus, extractor succeeds with overwhelming probability.

(Boneh-Shoup, Lemma 19.2)

If P^* succeeds with probability ε, then need to rely on "Rewinding Lemma" to argue that extractor obtains two accepting transcripts with probability at least $\varepsilon^2 - 1/p$.

The ability to extract a witness from any two accepting transcripts is very useful

↳ called special soundness (for 3-message protocols)

given (u, t_1, z_1) and (u, t_2, z_2) $\Rightarrow$ can extract the witness

initial message, challenge, response

[same initial message, different challenges]