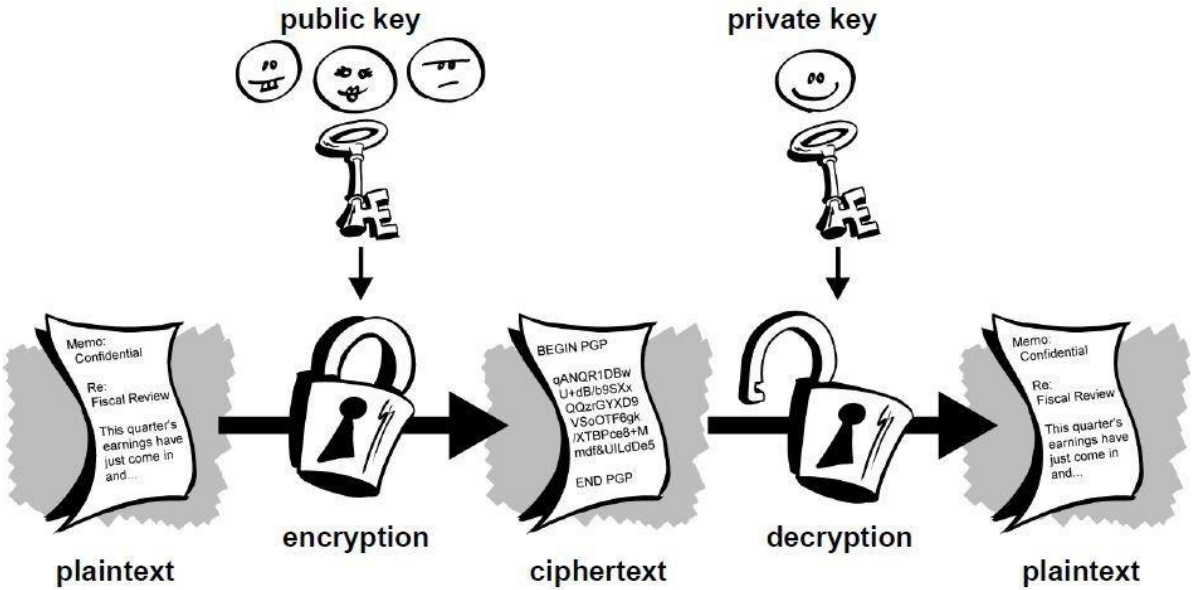
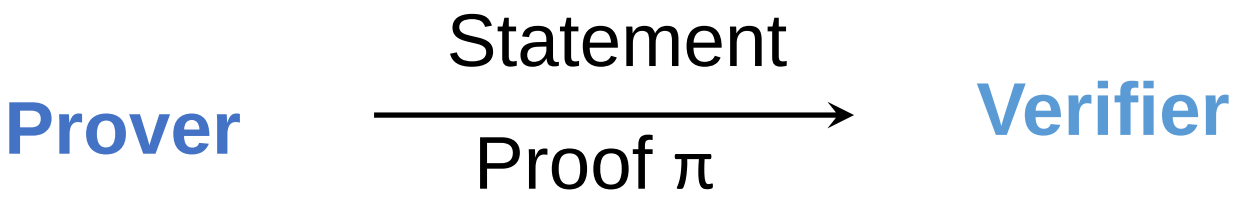
Zero-Knowledge Proofs

By
Vipul Goyal



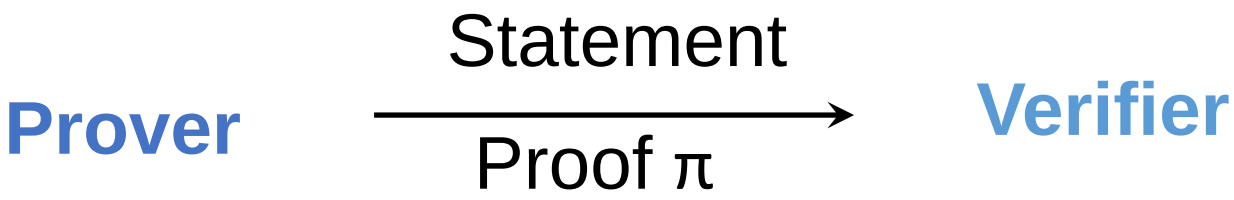
Traditional proofs



Example: $y^5 + y^2 + 7 = 0$

- Prover wants to prove that there exists value of y such that $y^5 + y^2 + 7 = 0$
- To prove: Prover can send this value to Verifier. Verifier can check equation satisfied.
- If prover lying: will be caught by Verifier

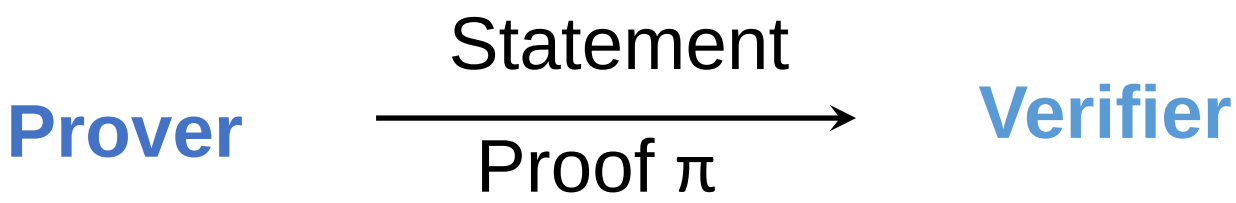
Traditional proofs



Another example: N is a product of two primes

- Prover wants to prove that N is a product of exactly two primes
- To prove: Prover can send these primes p, q to Verifier. Verifier verifies that p, q are primes and N = pq.
- If prover lying: will be caught by Verifier

Traditional proofs



Main Property:

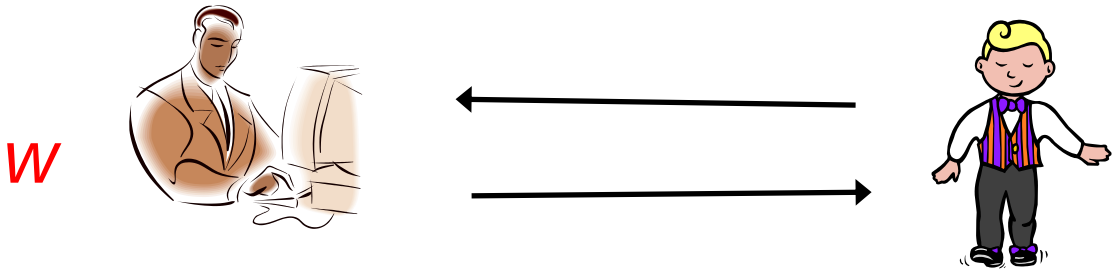
"**Soundness**"

If statement false, no matter what proof π* prover cooks up, verifier should detect the lie

Question: Can prover convince the verifier without revealing his secret? (Secret is also called witness.)

Zero Knowledge Proofs [GMR84]

Statement: $y^5 + y^2 + 7 = 0$



Prove an assertion (statement) without leaking anything

- Verifier convinced. Still doesn't learn the value of y.
- Is it even possible? How can verifier check without the right value of y?

Introduced by Goldwassar, Micali and Rackoff in 1982

Two Key Properties

"Zero-Knowledge"

Verifiers learns nothing "new" from the interaction except that the statement is true
(In particular, verifier doesn't learn prover's witness)

"**Soundness**"

If statement false, no matter what proof π^* prover cooks up, verifier should detect the lie with high probability

Note: to achieve ZK proofs, sometimes prover and verifier have to "talk" in multiple rounds

Meet a Magician

I can count the number of hair on your head

Really! Tell me?

3108312080

Meet a Magician

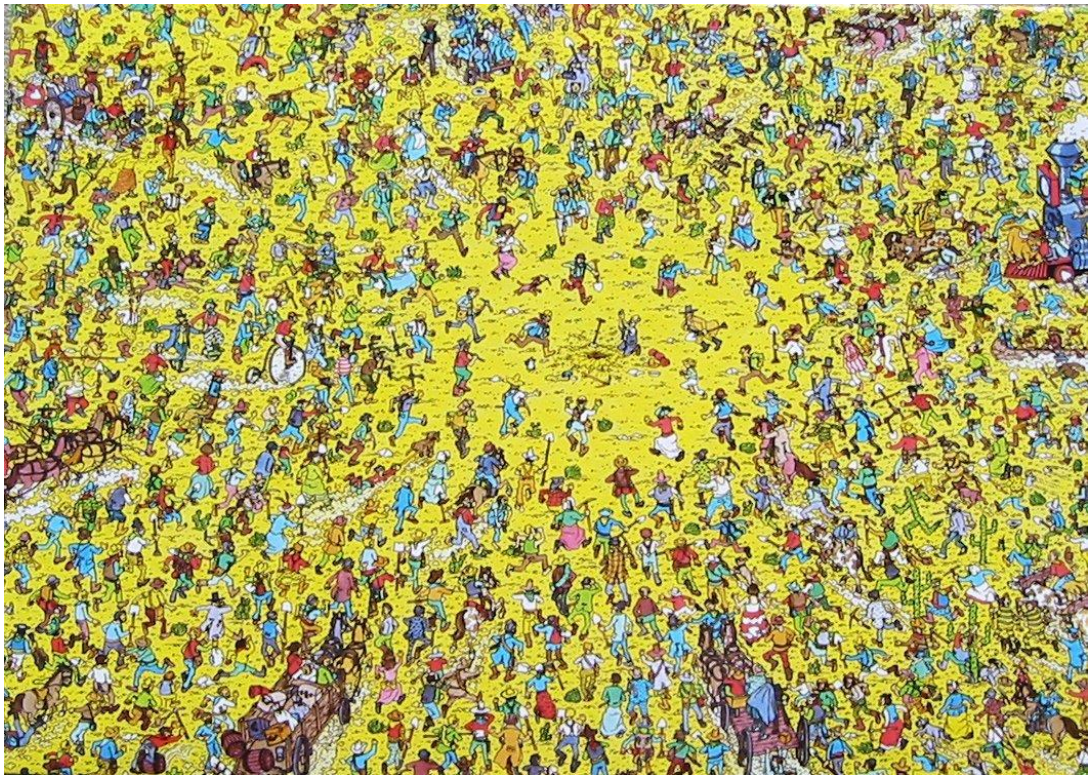
1) Go behind a door. Randomly take out either 0 or 1 hair. Come back and ask the Magician again.

2) If magician lying, can only succeed with probability ½

3) Repeat 100 times!

Where is Waldo?

- A famous British series of children puzzles
- Cover picture with paper. Just make a small hole to show Waldo.



Color-Blind Friend

- You have two balls: red and green. You have a color-blind friend.

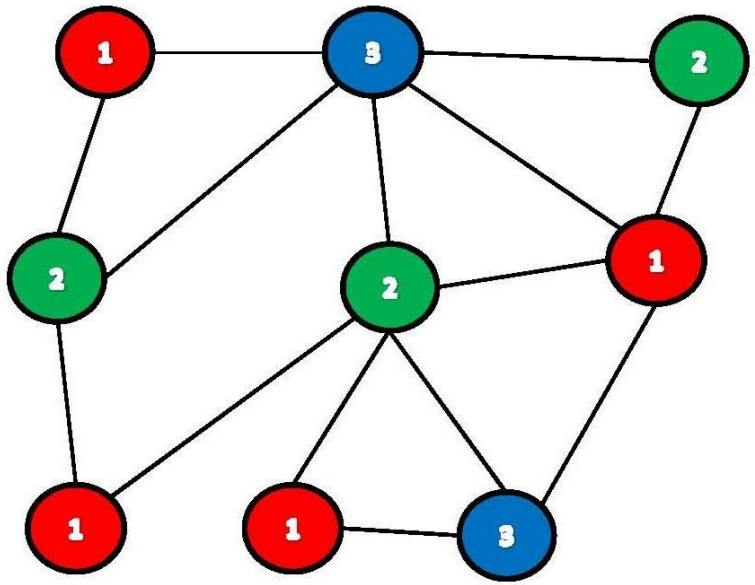
- You want to convince him that balls are different colors. But you don't want to reveal which ball is which color

1) Friend puts them behind his back. Randomly either exchanges or keeps the same
2) Shows you and you have to say "same" or "exchanged"
3) Repeat 100 times!

More Serious: Graph 3-Coloring Problem

Graph:

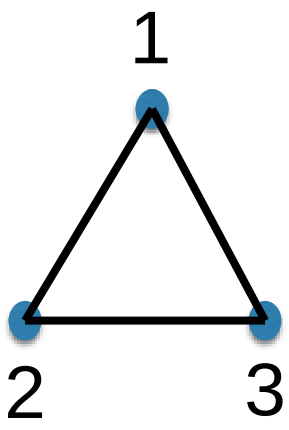
1) vertices (or nodes)
2) edges (connecting two vertices)

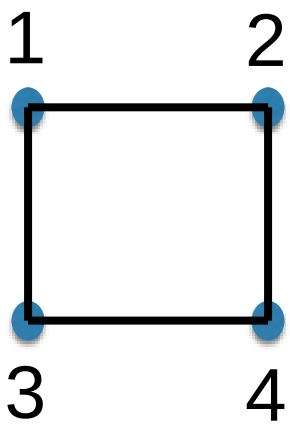


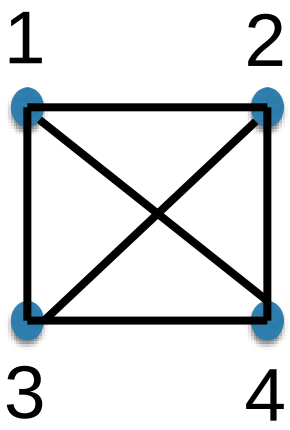
Graph 3 coloring: given a graph G, 3 colors (R, G, B), assign a color to every vertex

If two vertices share an edge, their colors should be different

Which graphs have 3 coloring?

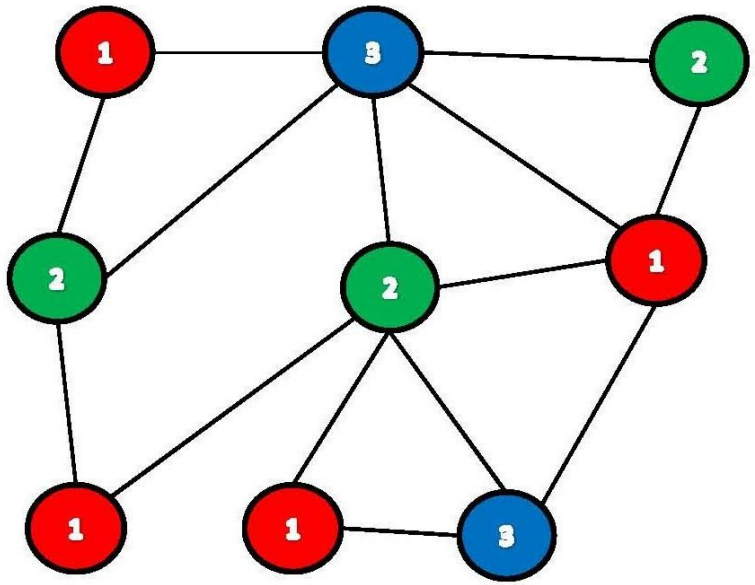






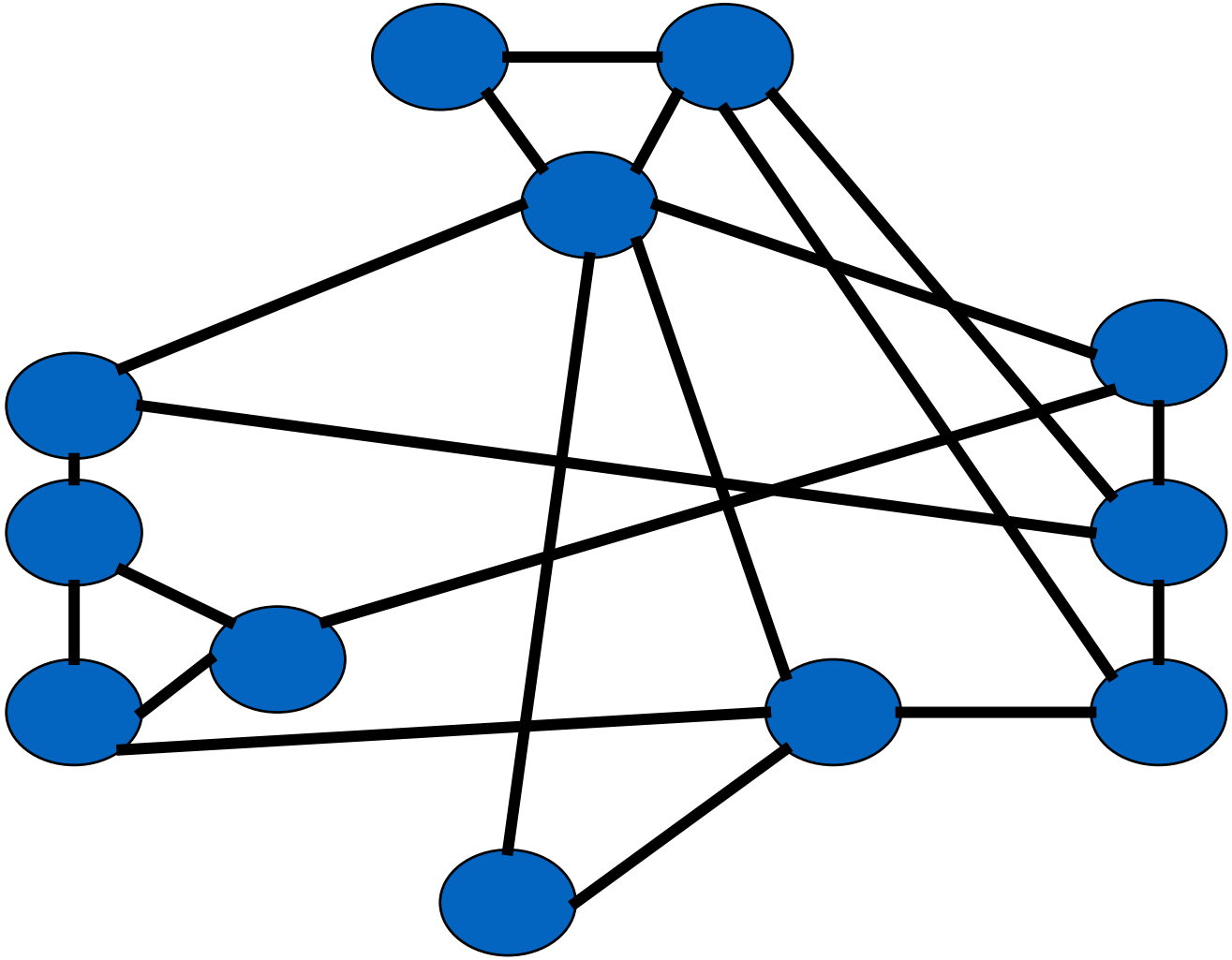
Why Graph 3-Coloring

Given: a graph G, 3 colors

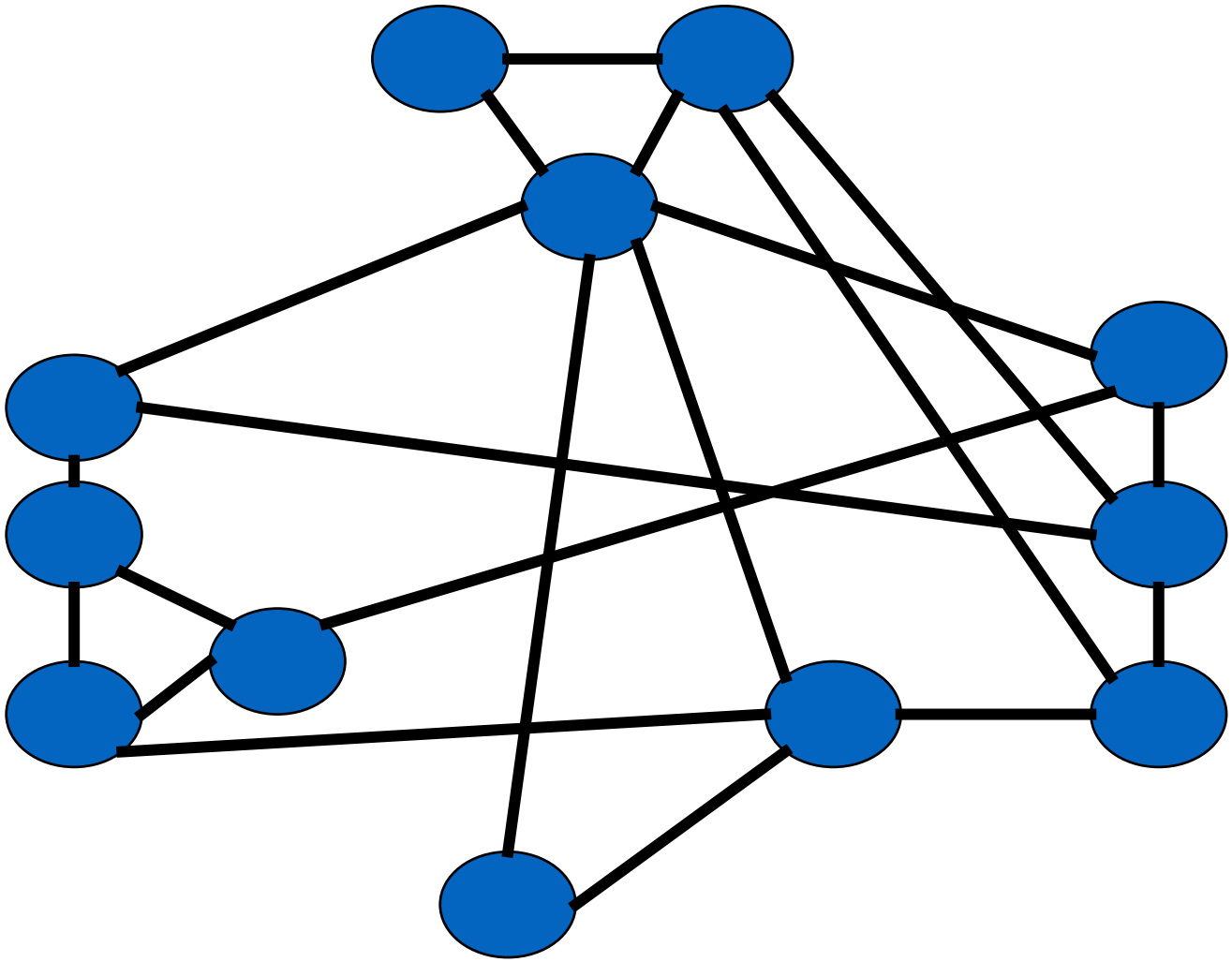


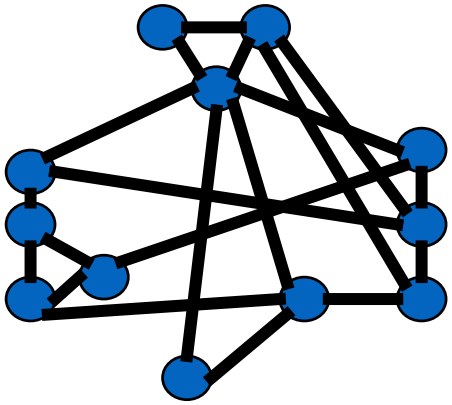
- Known to be an NP-complete problem (later)
- Very hard to solve: no good algorithms to color vertices
- Almost every algorithmic problem of interest can be reduced to graph 3-coloring (e.g. factoring or solving discrete log)

Prover and verifier have a graph
Prover knows a 3 coloring
He can send it to verifier: verifier checks every edge



But this reveals the entire coloring to verifier!
Can Prover give a “Zero Knowledge” proof of this fact?



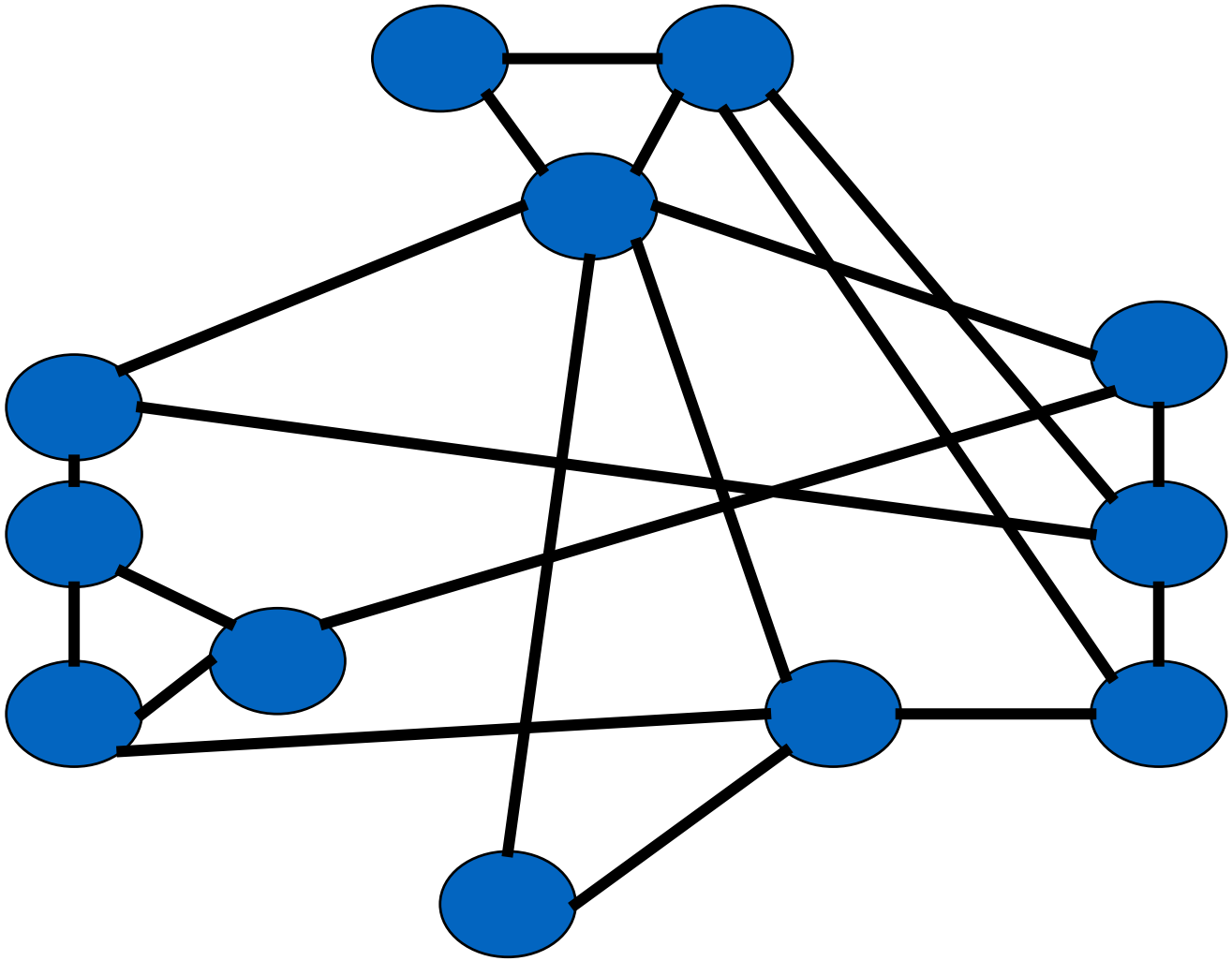


Several ways to permute the color names

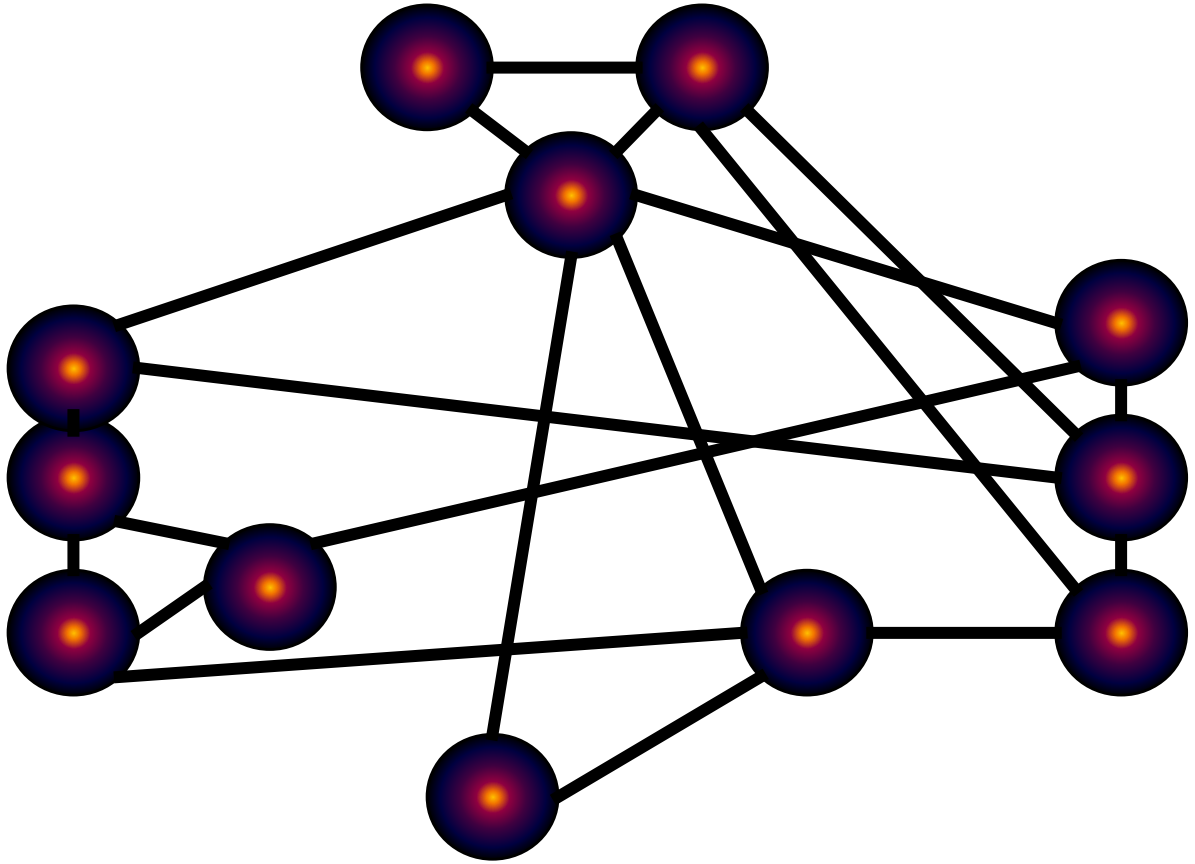
Example: if I exchange R and B, the coloring still remains a valid 3-coloring

Overall, given a 3-coloring of graph, there are 6 colorings that can be obtained by permuting the 3 color names (3! colorings)

Prover randomly chooses one of the 6 colorings obtainable from the secret coloring.

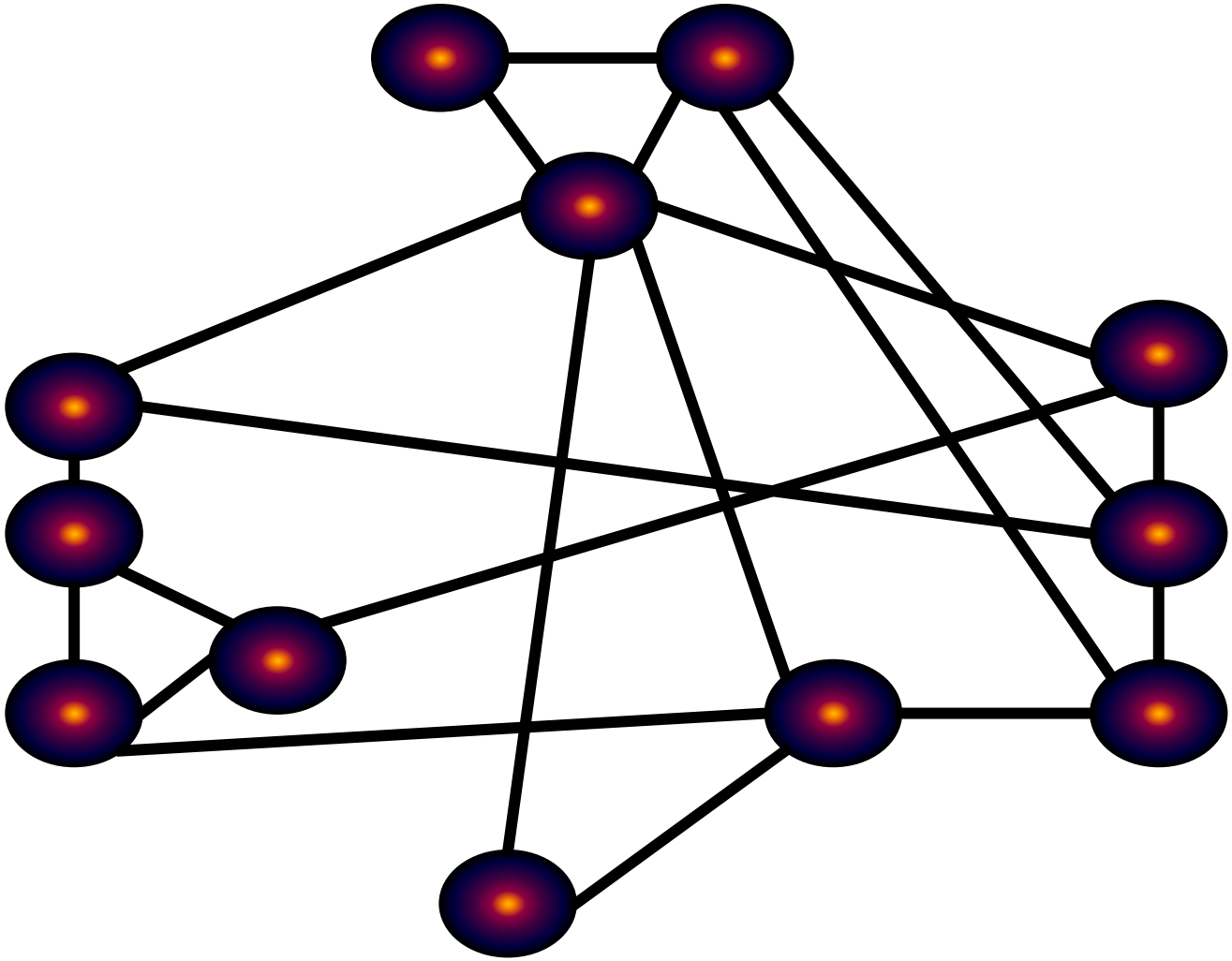


Prover randomly chooses one of the 6 colorings obtainable from the secret coloring.

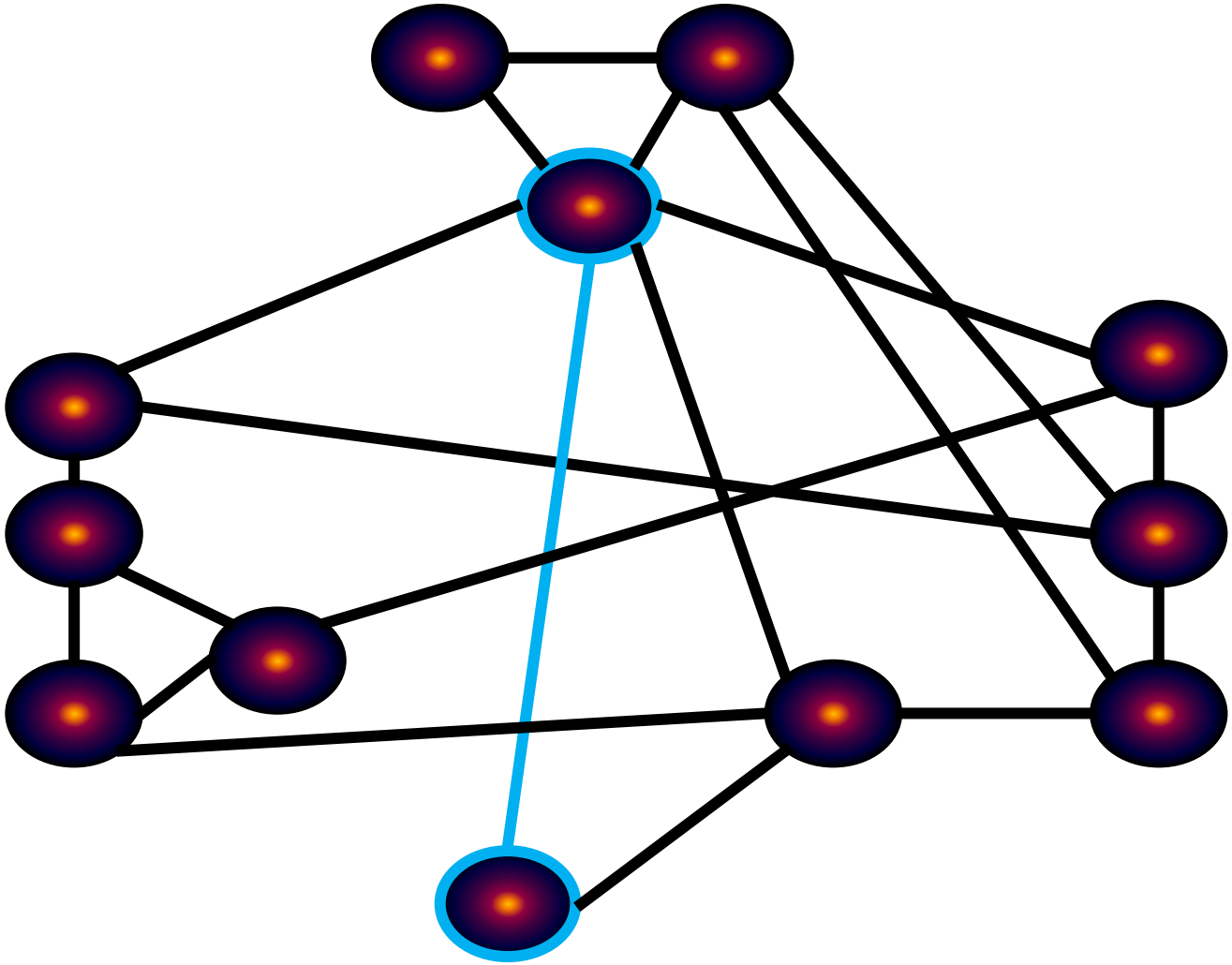


Prover and verifier in the same room
Each node will have an associated “locked box.”
The prover places the colors in the corresponding boxes

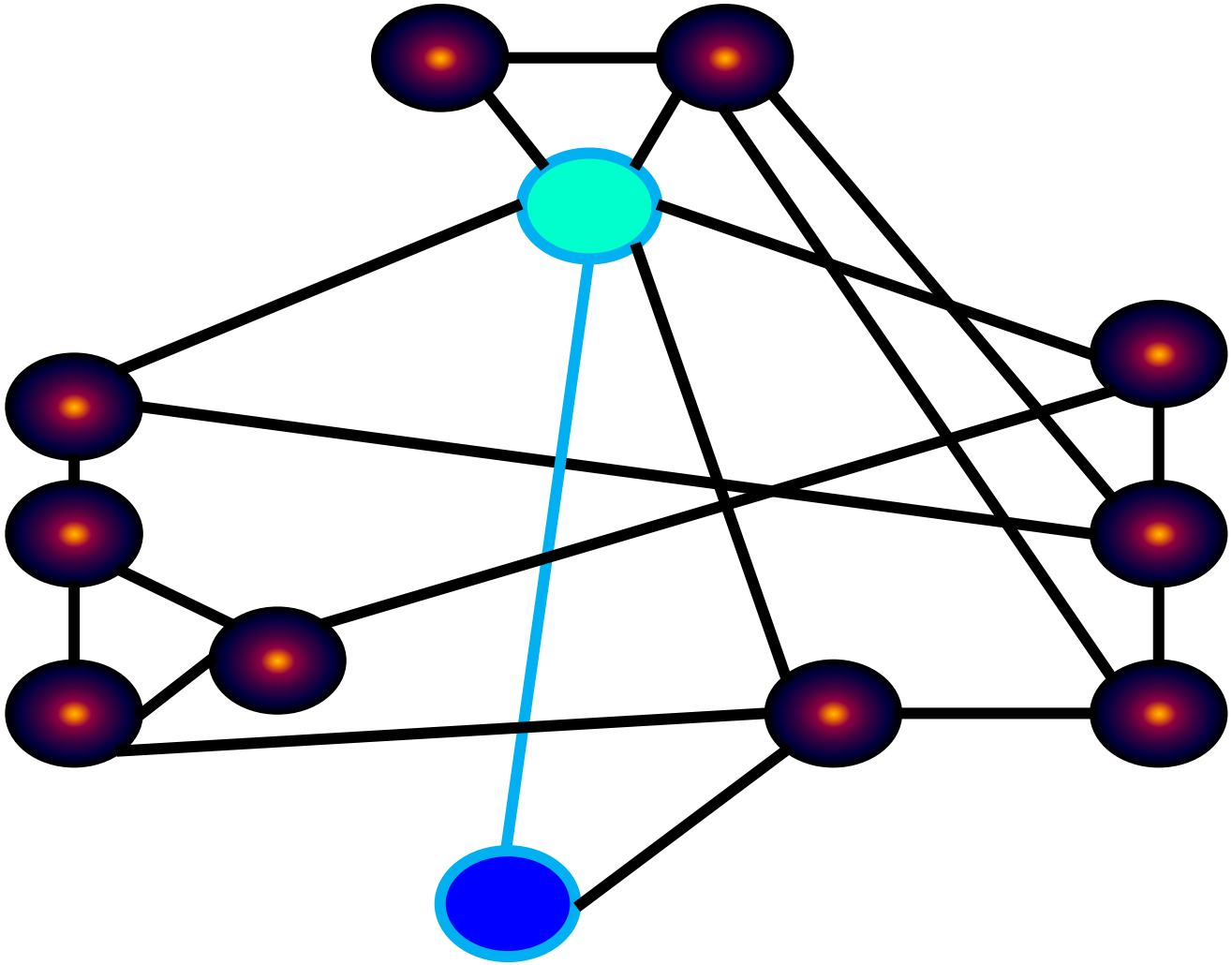
Verifier: Picks an edge at random and asks to open the boxes at both ends of the edge.



Verifier: Picks an edge at random and asks to open the boxes at both ends of the edge.

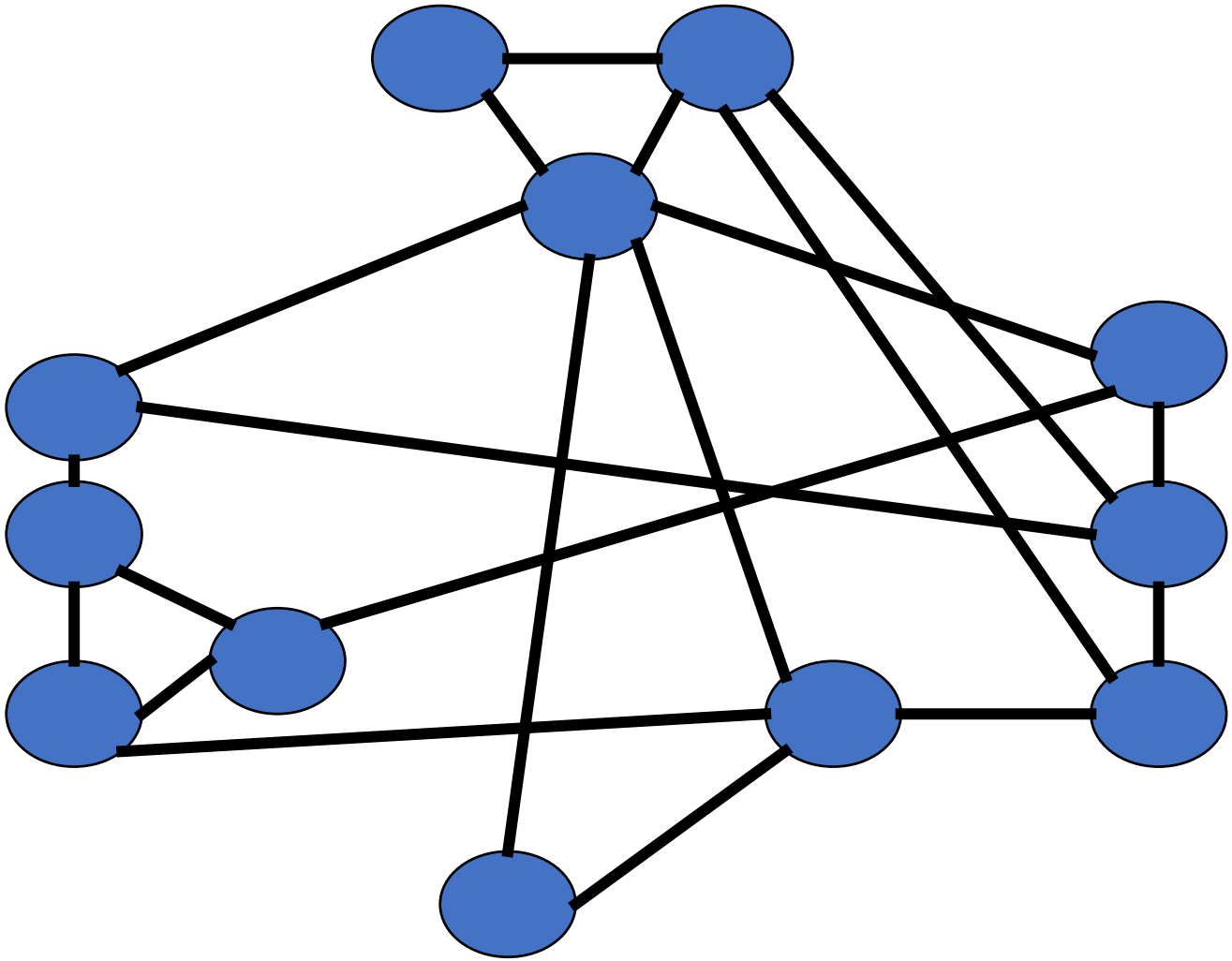


Verifier: Picks an edge at random and asks to open the boxes at both ends of the edge.

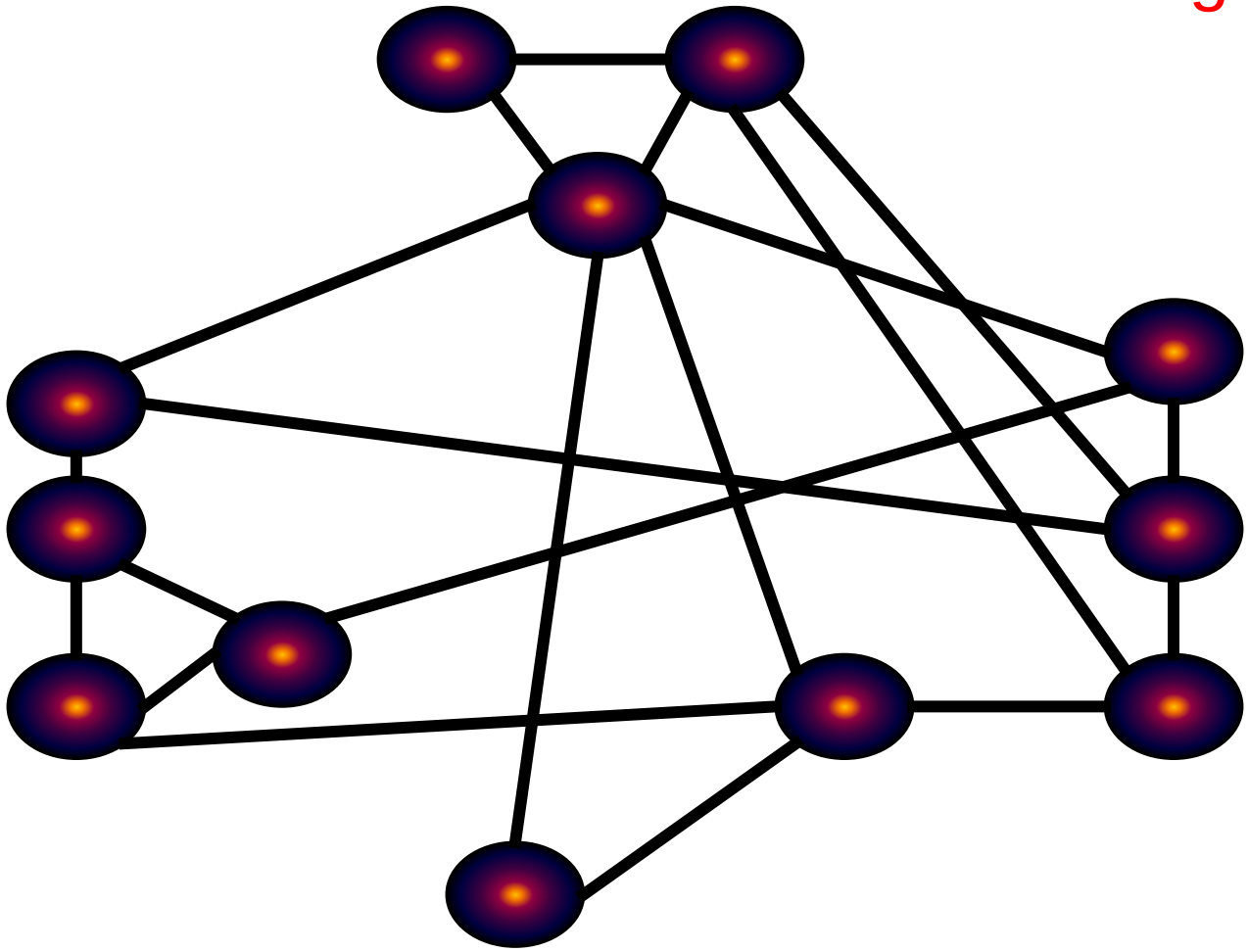


REPEAT USING AN INDEPENDENT, RANDOM CHOICE OF THE 6 PERMUTATIONS OF COLOR NAMES

Prover randomly chooses one of the 6 colorings obtainable from the secret coloring.

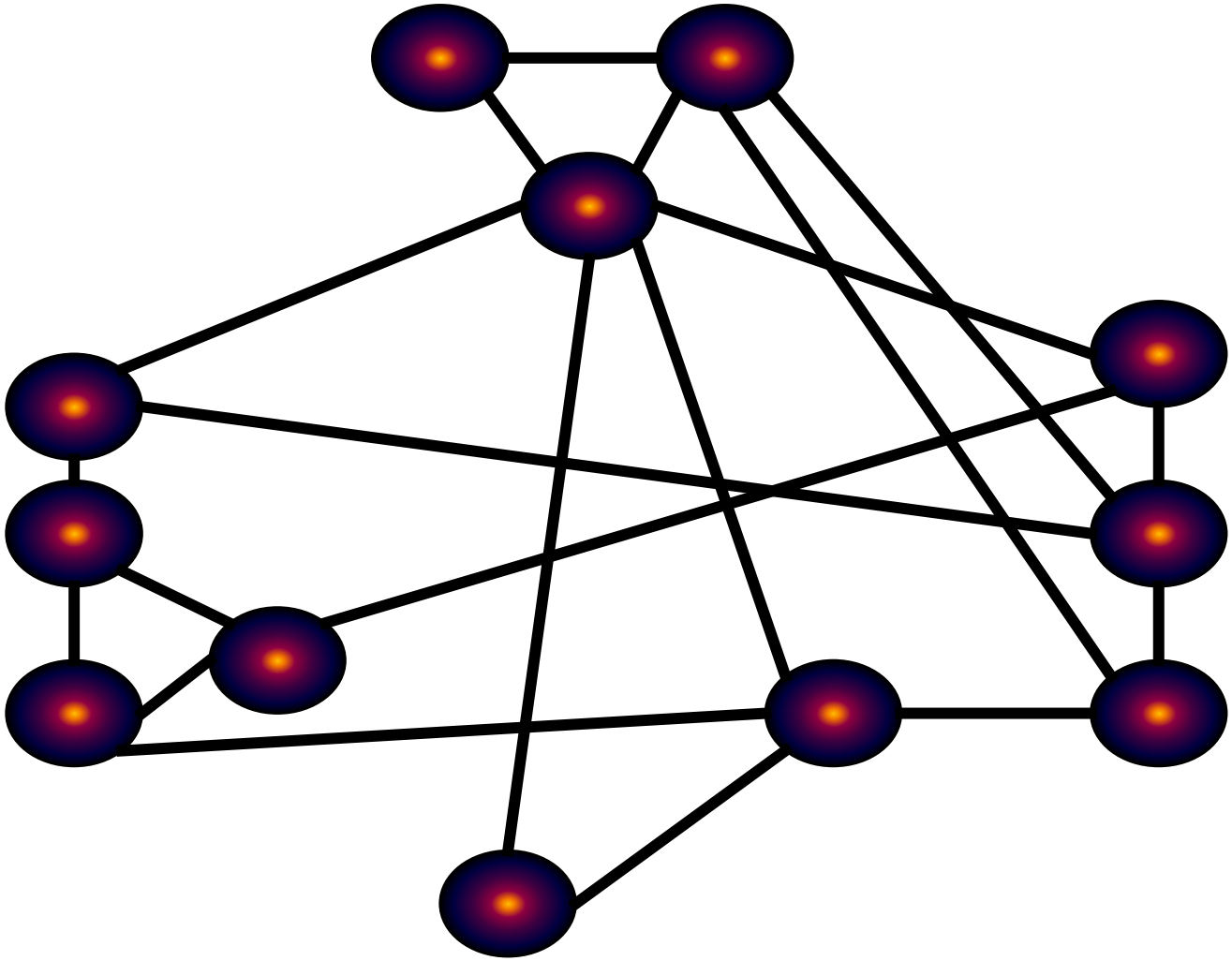


Prover randomly chooses one of the 6 colorings obtainable from the secret coloring.

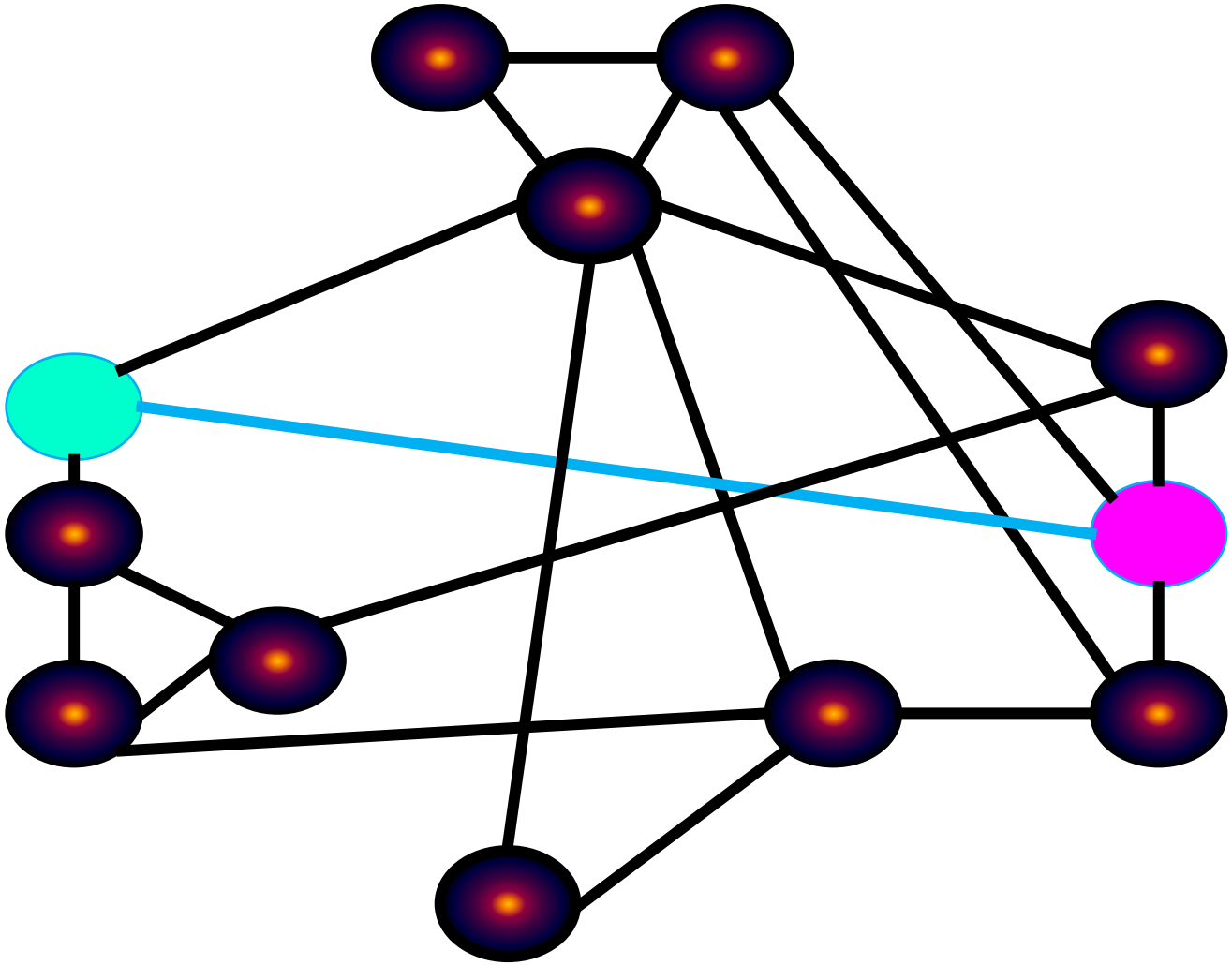


Each node will have an associated “locked box.”
The prover places the colors in the corresponding boxes

Verifier: Picks an edge at random and asks to open the boxes at both ends of the edge.



Verifier: Picks an edge at random and asks to open the boxes at both ends of the edge.



Prover repeats with independent permutations

If *all* iterations accept, verifier accepts

Is it a Valid ZK Proof?

Soundness: If graph not 3-colorable, at least one edge has same color. 1/m chance of catching.

Repeat cm times

Probability of fooling the Verifier for all cm challenges $= (1 - 1/m)^{cm} < (1/e)^{c}$

Taking c=100, this is very small

Is it a Valid ZK Proof?

Zero-Knowledge

In each iteration, Verifier sees two different random colors:

- I.e., one of (1,2),(2,3),(3,1), (2,1),(3,2),(1,3) picked at random
- Nothing he didn't know/anticipate already!!

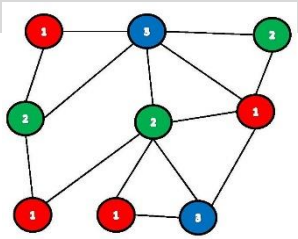
Locked Boxes?

What if Prover and verifier are not in the same room?

Commitment Schemes

- Commitment like a note placed in a safe
- Two properties: hiding and binding
- Electronic equivalent of such a safe

ZK for Graph 3-Coloring



Randomly permute

Prover

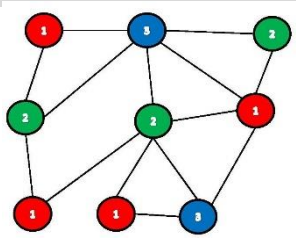
$\text{Com}(V_1, \text{red}), \ldots, \text{Com}(V_n, \text{blue})$ →

← $e_{ij} = (V_i, V_j)$

Open $\text{Com}(V_i, \text{red}), \text{Com}(V_j, \text{green})$ →

Verifier

What Did we Achieve?



Prover

Verifier

- Zero-Knowledge for graph 3-coloring problem
- Hundreds of interesting languages?
- Exploit NP completeness

P, NP and NP-Completeness

- P = class of problems which you can solve efficiently on your own (in polynomial time). Think: multiplying two numbers.
- NP = class of problems whose solutions can be verified efficiently (in poly time). Thinking: factoring a given number. P is a subset of NP.
- NP-completeness: any problem in NP can be converted into an instance of an NP-complete problem! Solve this, solve original problem!
- P = NP: if you can verify a problem efficiently, you can also solve it efficiently!

P vs NP Problem

CMI

ABOUT PROGRAMS MILLENNIUM PROBLEMS PEOPLE PUBLICATIONS EUCLID EVENTS

Millennium Problems

Yang–Mills and Mass Gap

Experiment and computer simulations suggest the existence of a "mass gap" in the solution to the quantum versions of the Yang-Mills equations. But no proof of this property is known.

1 million dollar question

Riemann Hypothesis

The prime number theorem determines the average distribution of the primes. The Riemann hypothesis tells us about the deviation from the average. Formulated in Riemann's 1859 paper, it asserts that all the 'non-obvious' zeros of the zeta function are complex numbers with real part 1/2.

P vs NP Problem

If it is easy to check that a solution to a problem is correct, is it also easy to solve the problem? This is the essence of the P vs NP question. Typical of the NP problems is that of the Hamiltonian Path Problem: given N cities to visit, how can one do this without visiting a city twice? If you give me a solution, I can easily check that it is correct. But I cannot so easily find a solution.

Navier–Stokes Equation

This is the equation which governs the flow of fluids such as water and air. However, there is no proof for the most basic questions one can ask: do solutions exist, and are they unique? Why ask for a proof? Because a proof gives not only certitude, but also understanding.

Hodge Conjecture

The answer to this conjecture determines how much of the topology of the solution set of a system of algebraic equations can be defined in terms of further algebraic equations. The Hodge conjecture is known in certain special cases, e.g., when the solution set has dimension less than four. But in dimension four it is unknown.

Poincaré Conjecture

In 1904 the French mathematician Henri Poincaré asked if the three dimensional sphere is characterized as the unique simply connected three manifold. This question, the Poincaré conjecture, was a special case of Thurston's geometrization conjecture. Perelman's proof tells us that every three manifold is built from a set of standard pieces, each with one of eight well-understood geometries.

Birch and Swinnerton-Dyer Conjecture

Supported by much experimental evidence, this conjecture relates the number of points on an elliptic curve mod p to the rank of the group of rational points. Elliptic curves, defined by cubic equations in two variables, are fundamental mathematical objects that arise in many areas: Wiles' proof of the Fermat Conjecture, factorization of numbers into primes, and cryptography, to name three.

ZK for all of NP

$y^5 + y^2 + 7 = 0$

w

Prover

$y^5 + y^2 + 7 = 0$

Verifier

NP Reduction

Graph G

w'

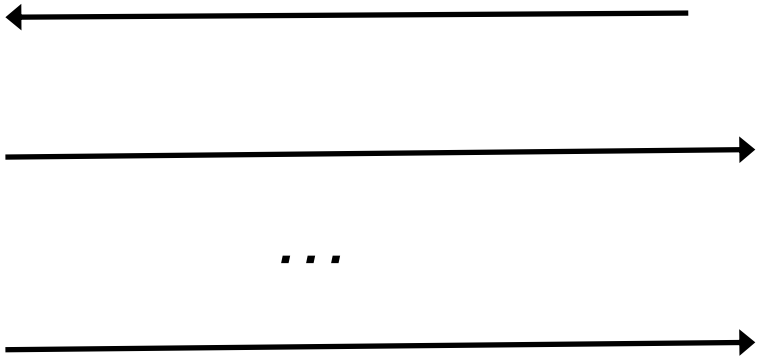
Graph G

A Beautiful Fact

Everything provable is provable
in zero-knowledge!

Some Other Cool Primitives

Secure 2-party Computation



x

y

f(x,y)

- Yao's Millionaire Problem: who has more money?
- Secure 2pc: Learn only f(x,y) but not x or y
- Can extend to multi-party

MPC Applications

- Privacy preserving machine learning
- Auctions / Voting
- Storing and dealing with private data on blockchain
- ……..

A Simple MPC Protocol

- Let's compute the average age of a student in our class
  - Assume nobody will lie about their age. Also, a student can whisper in the ears of the next student.
  - I start with a large random number (say a few millions) R. Whisper it to the first student P_1.
  - P_1 computes $R+A_1$ and whispers to P_2, … and so on
  - P_n whispers back $R+A_1+A_2+….A_n$ to me
  - I subtract R, divide by n and announce the result.

- Look at any student: they just receive some random number and nothing else
- I receive $A_1+A_2+….A_n$ and nothing else

MPC Results

- What about functions other than sum/average?
- What if parties collude with each other?
- Known: MPC protocols for any function f for any number of parties.
- Parties can be dishonest and collude with each other. Still can’t learn extra information about the input of an honest party!

Fully Homomorphic Encryption
[Gentry'09]

Sure. It will be $50.

Need expensive DNA sequencing computation done

Send me the Data

Ok. Sounds good.

Sorry I can't. Its too sensitive

Enc(x), f

Enc(f(x))

Questions?