Cryptographic hash functions II

Pollard rho for collision search

Tanja Lange

Eindhoven University of Technology

2MMC10 – Cryptology

Generic hardness

The birthday paradox implies that if one draws elements at random from a set of m elements, then with 50% probability one has picked one element twice after about $\sqrt{\pi m/2}$ picks.
Hence it takes $O(2^{n/2})$ calls to H to find a collision.

Generic hardness

The birthday paradox implies that if one draws elements at random from a set of m elements, then with 50% probability one has picked one element twice after about $\sqrt{\pi m/2}$ picks.
Hence it takes $O(2^{n/2})$ calls to H to find a collision.

Two problems need to be solved to use the rho method successfully:

1. Design step function so that it "randomly" samples elements (so that the birthday paradox applies) while being deterministic (so we can use Floyd's cycle finding method to remove storage).
2. Design step function so that collision gives meaningful result.

Generic hardness

The birthday paradox implies that if one draws elements at random from a set of m elements, then with 50% probability one has picked one element twice after about $\sqrt{\pi m/2}$ picks.
Hence it takes $O(2^{n/2})$ calls to H to find a collision.

Two problems need to be solved to use the rho method successfully:

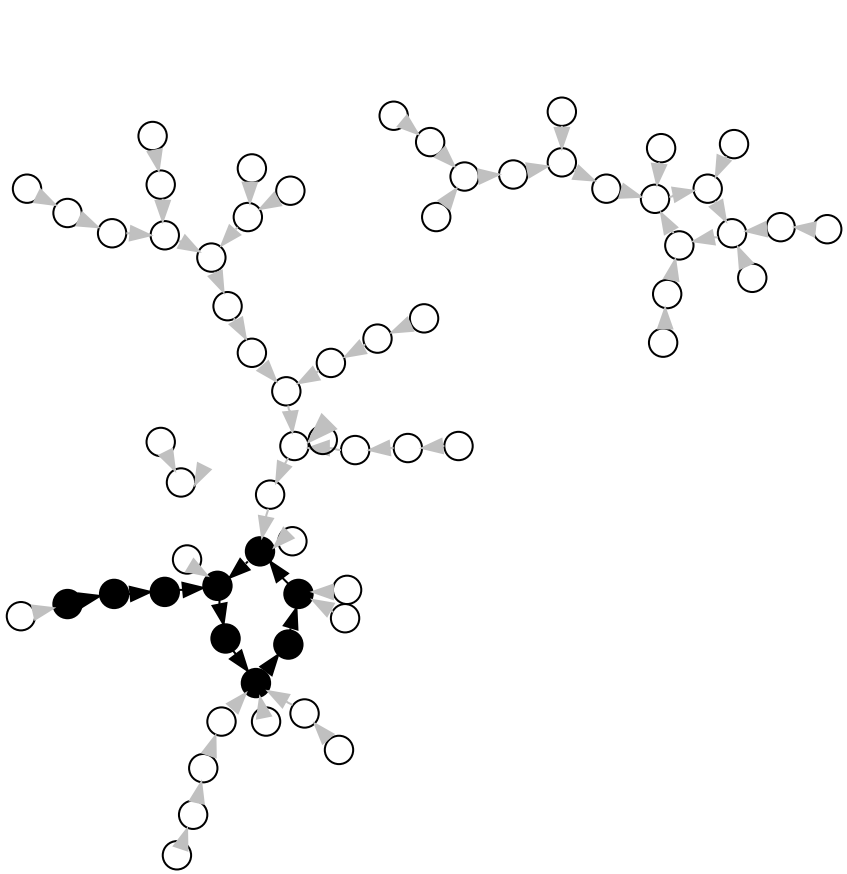
1. Design step function so that it "randomly" samples elements (so that the birthday paradox applies) while being deterministic (so we can use Floyd's cycle finding method to remove storage).
2. Design step function so that collision gives meaningful result.

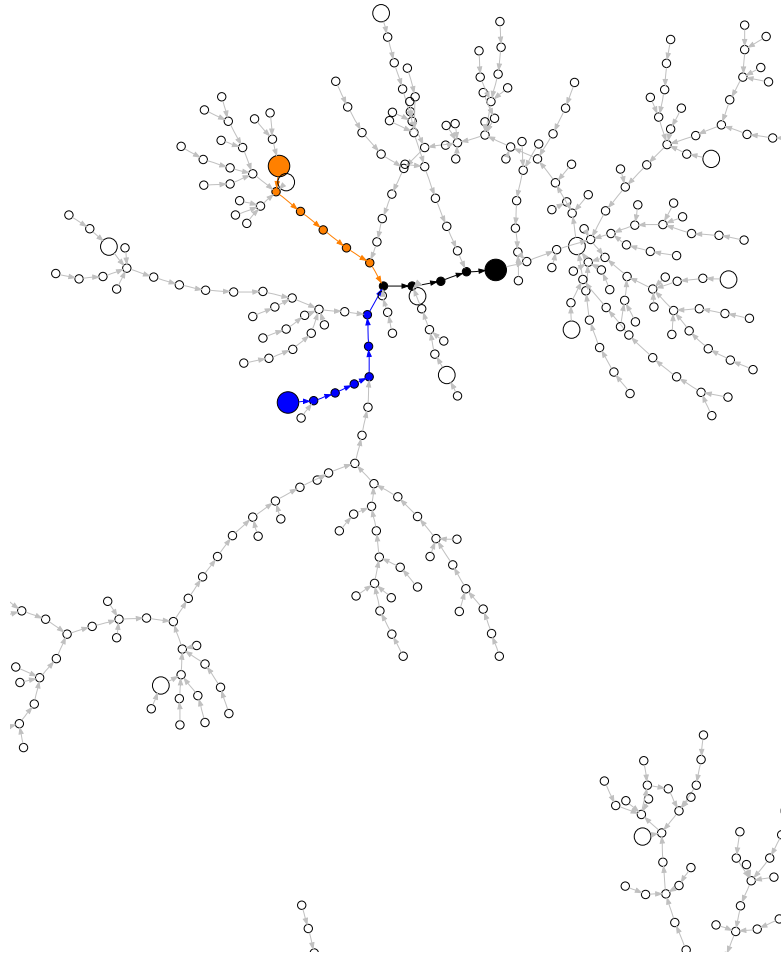
If message space includes output space, e.g., $H : \{0,1\}^* \to \{0,1\}^n$, use random starting point W_0 and iterate $W_{i+1} = H(W_i)$.

Else $H : \mathbf{M} \to \mathbf{H}$ for some message space $\mathbf{M}$ and hash space $\mathbf{H}$, compose with some map $\phi : \mathbf{H} \to \mathbf{M}$, iterate $W_{i+1} = H(\phi(W_i))$.

Can use parallel version due to van Oorshot and Wiener.

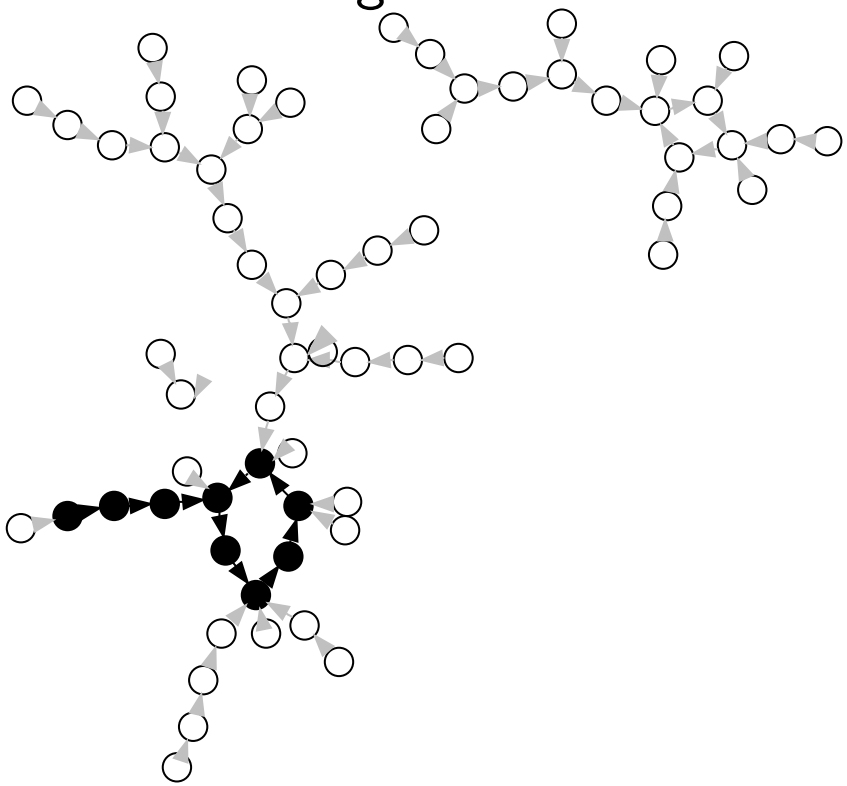
How to use collisions?

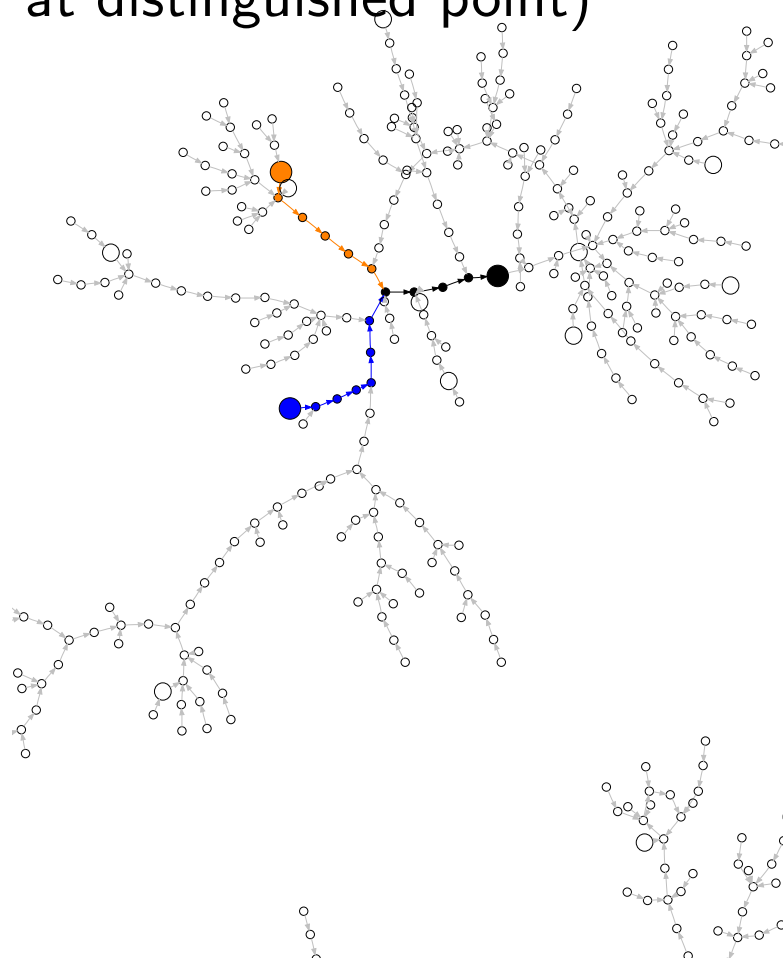




How to use collisions?

When the collision is detected (in Floyd or at distinguished point) it is not a meaningful collision!

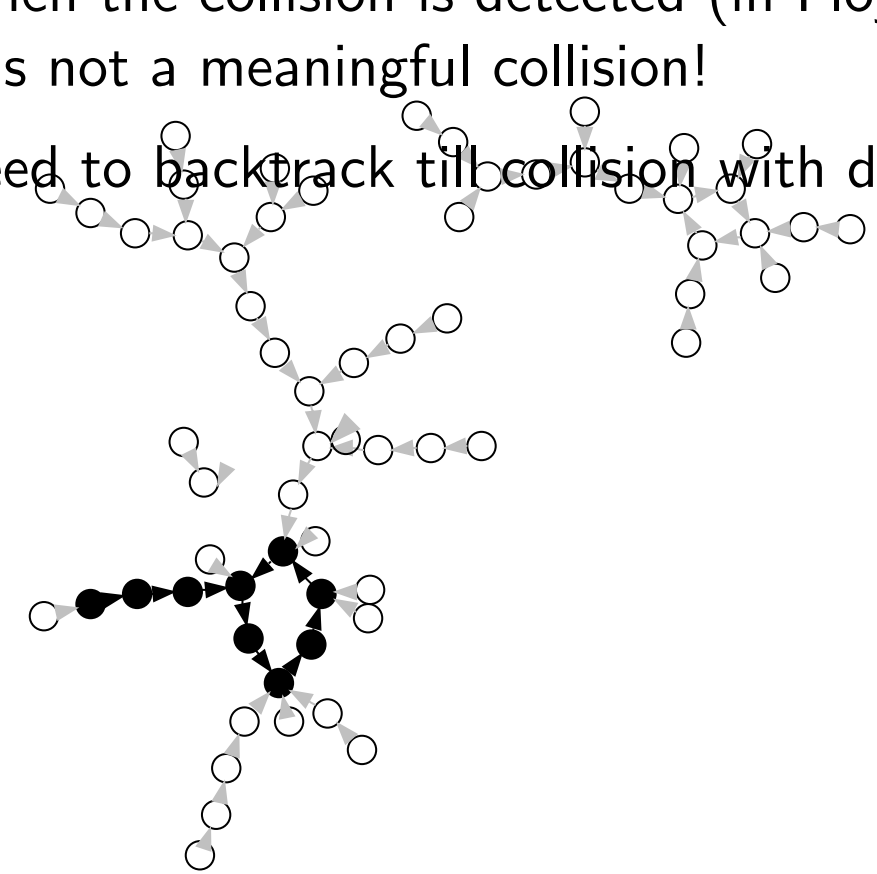


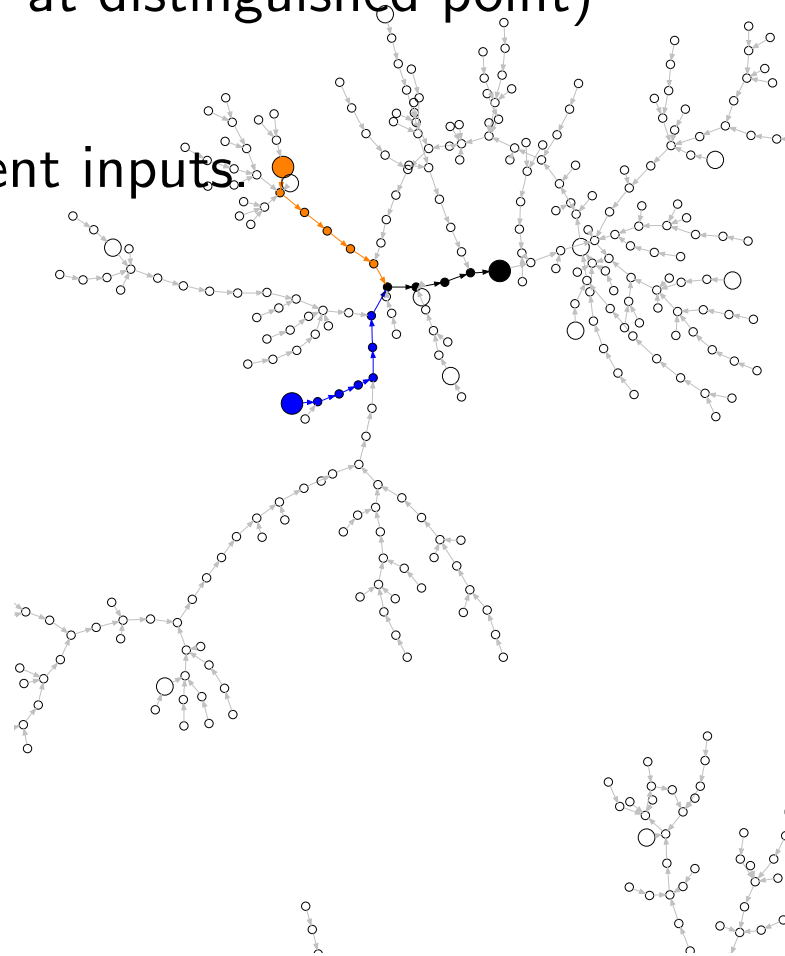


How to use collisions?

When the collision is detected (in Floyd or at distinguished point) it is not a meaningful collision!

Need to backtrack till collision with different inputs.

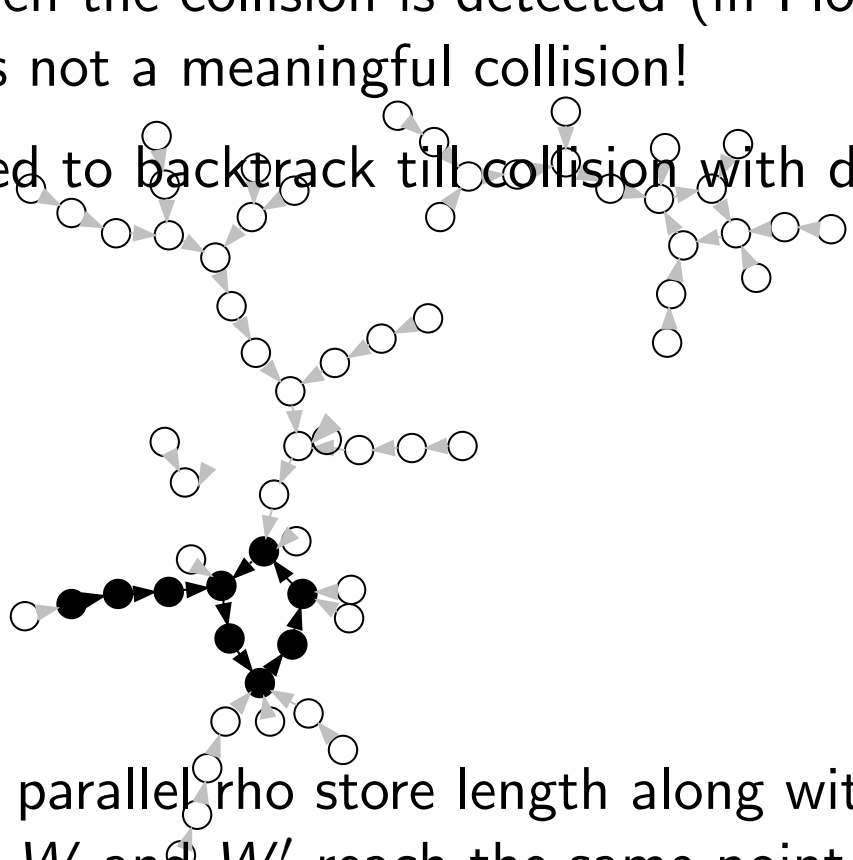


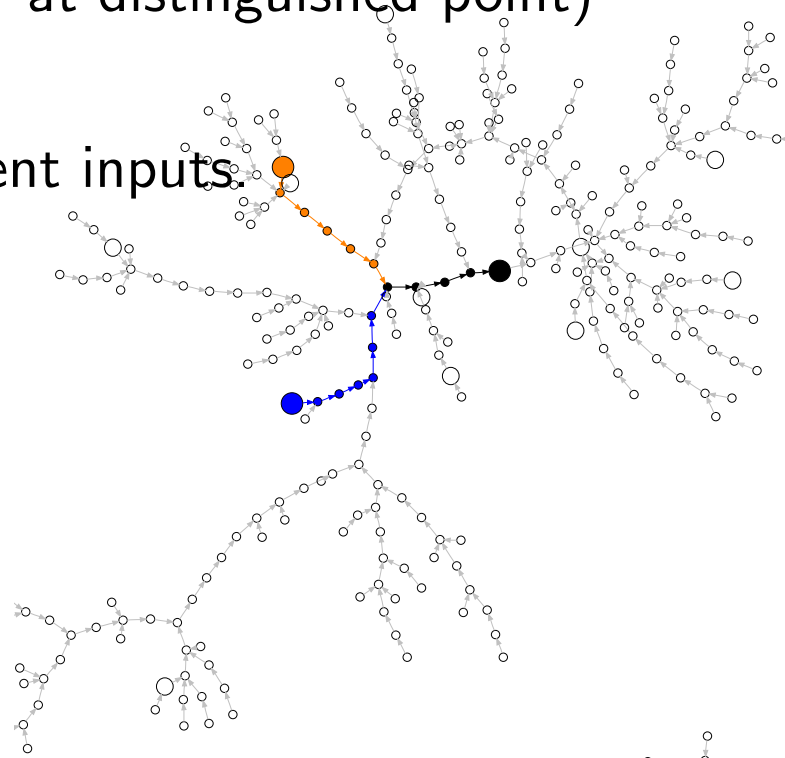


How to use collisions?

When the collision is detected (in Floyd or at distinguished point) it is not a meaningful collision!

Need to backtrack till collision with different inputs.





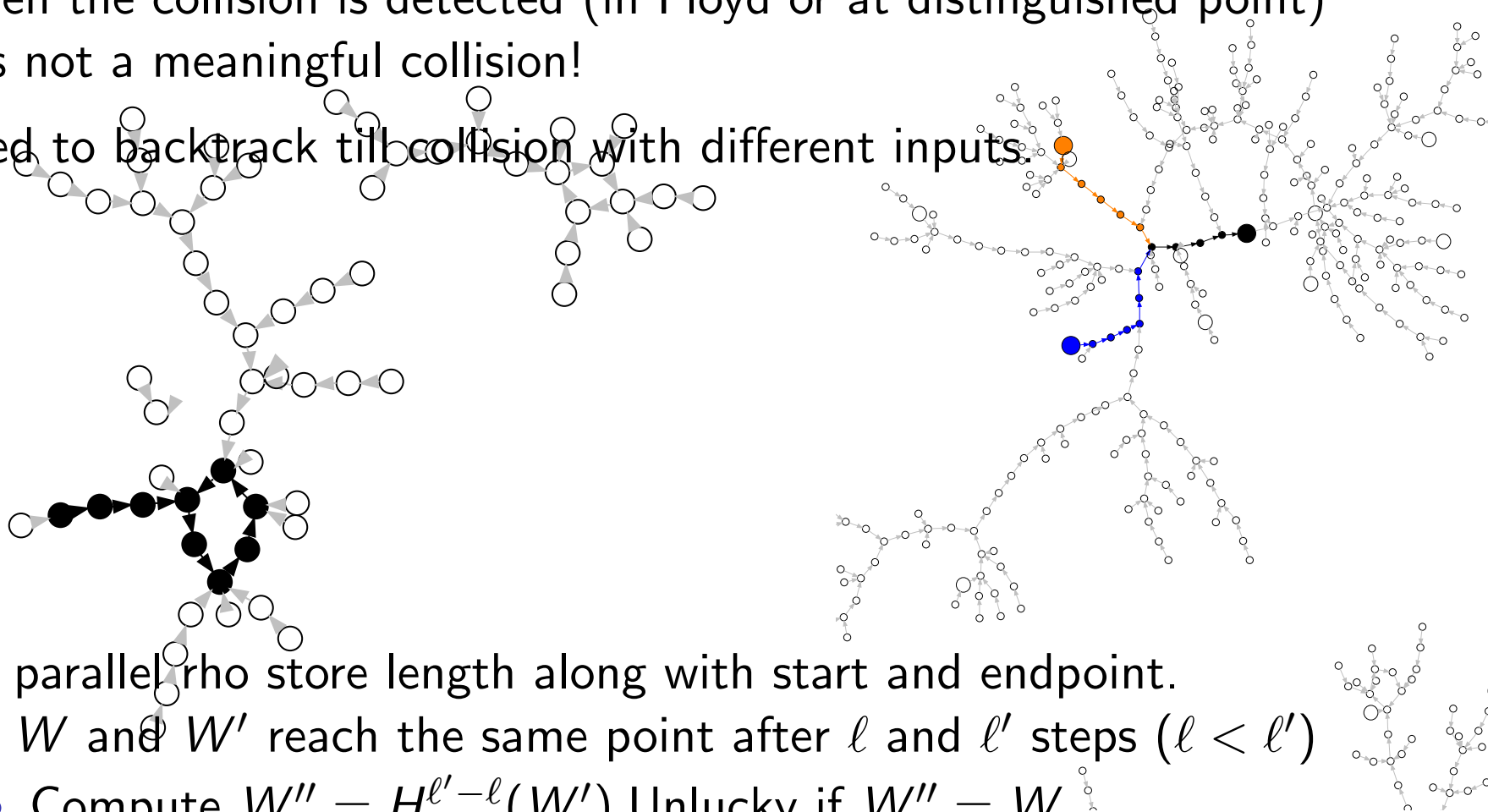
For parallel rho store length along with start and endpoint.
Let W and W' reach the same point after ℓ and ℓ' steps ($\ell < \ell'$)

- Compute $W'' = H^{\ell'-\ell}(W')$.

How to use collisions?

When the collision is detected (in Floyd or at distinguished point) it is not a meaningful collision!

Need to backtrack till collision with different inputs.



For parallel rho store length along with start and endpoint.

Let W and W' reach the same point after ℓ and ℓ' steps ($\ell < \ell'$)

- Compute $W'' = H^{\ell'-\ell}(W')$.Unlucky if $W'' = W$.
- Iterate the following three steps until a collision is found:
  - Compare W and W''. If they are equal, output W_o and W_o''.
  - Store $W_o = W, W_o'' = W''$.
  - Update $W = H(W), W'' = H(W'')$.

How to get interesting collisions?

We typically want meaningful collisions, e.g., colliding pdf files. This is one of the places where $\mathbf{H} \not\subset \mathbf{M}$ happens.

Specify n bit positions in $\mathbf{M}$ that permit variations, make ϕ map to those injectively. Several advanced formats are very flexible in taking macros or invisible images.

How to get interesting collisions?

We typically want meaningful collisions, e.g., colliding pdf files. This is one of the places where $\mathbf{H} \not\subset \mathbf{M}$ happens.

Specify n bit positions in $\mathbf{M}$ that permit variations, make ϕ map to those injectively. Several advanced formats are very flexible in taking macros or invisible images.

Even more meaningful collisions

- Chosen prefix collision: want $m = p*$ and $m' = p'*$ to collide, where $*$ stands for arbitrary strings.
- Read Predicting the winner of the 2008 US Presidential Elections using a Sony PlayStation 3 for an impressive example of 12 pdf files all having the same hash, using chosen prefix collisions in MD5 and the flexibility of the pdf standard.