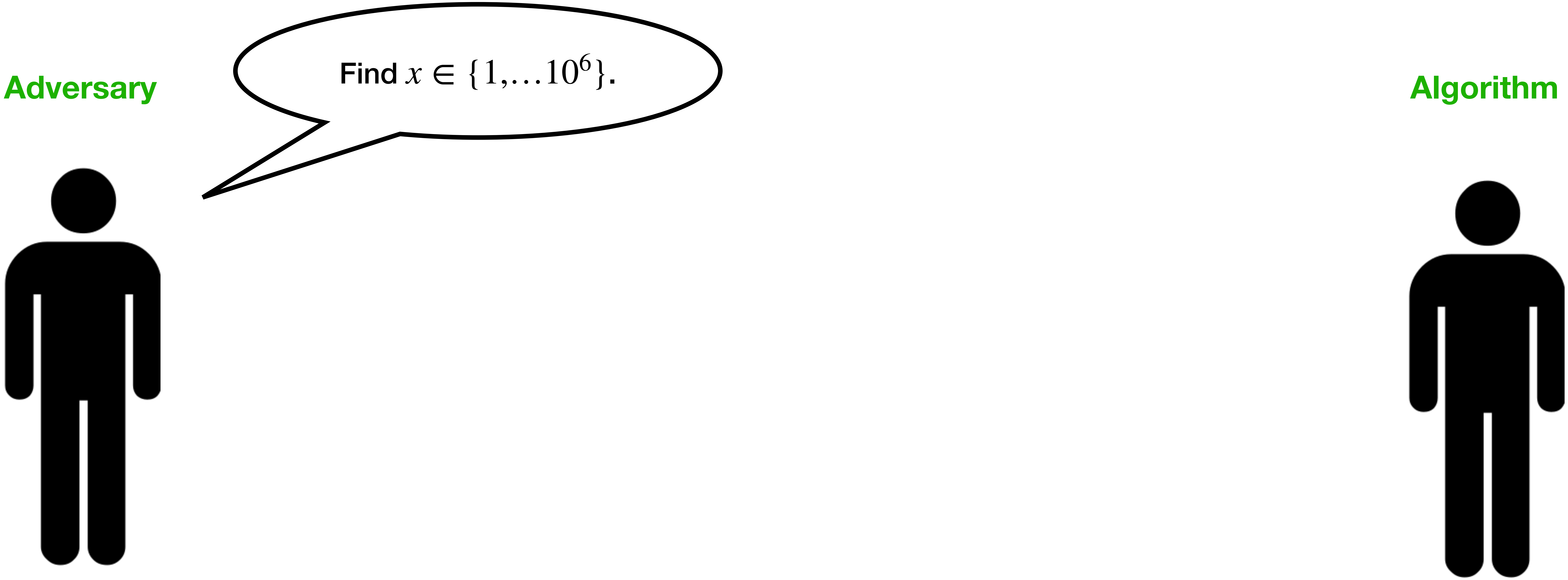
Noisy (Binary) Searching: Simple, Fast and Correct

By Dariusz Dereniowski, Aleksander Łukasiewicz and Przemysław Uznański

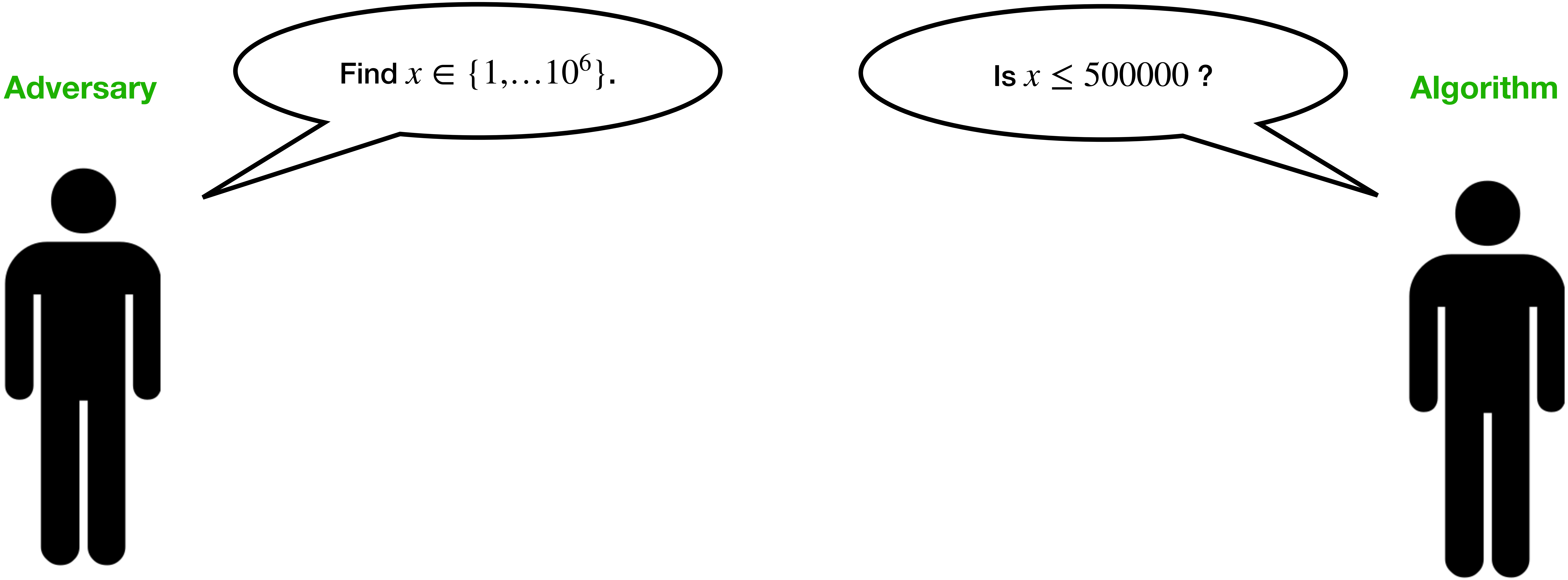
Part I

The problem, the model and all that jazz…

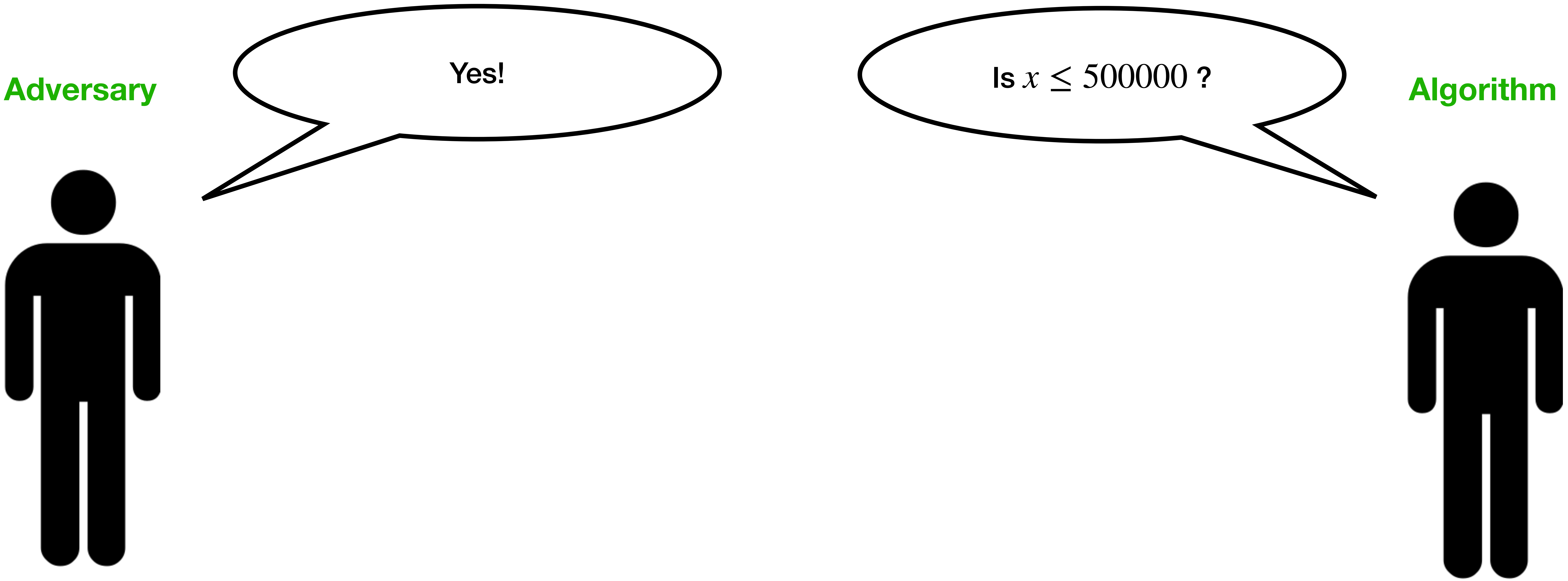
Binary search as a game



Binary search as a game

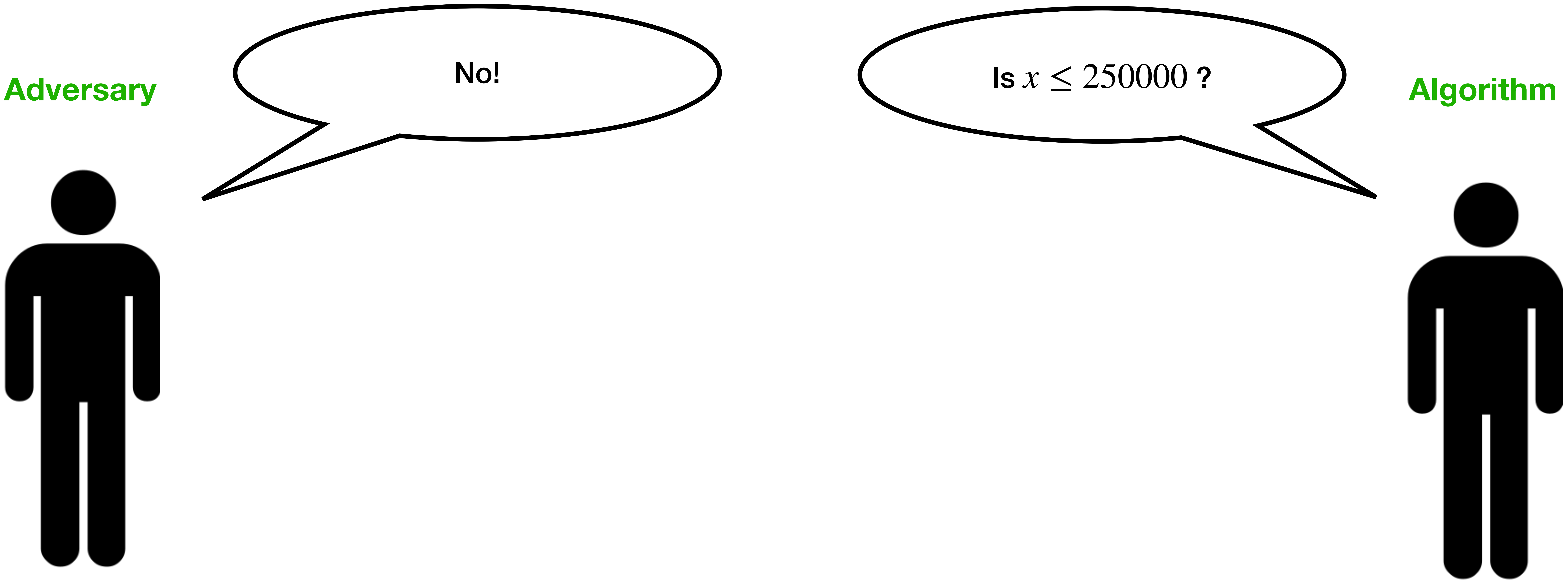


Binary search as a game



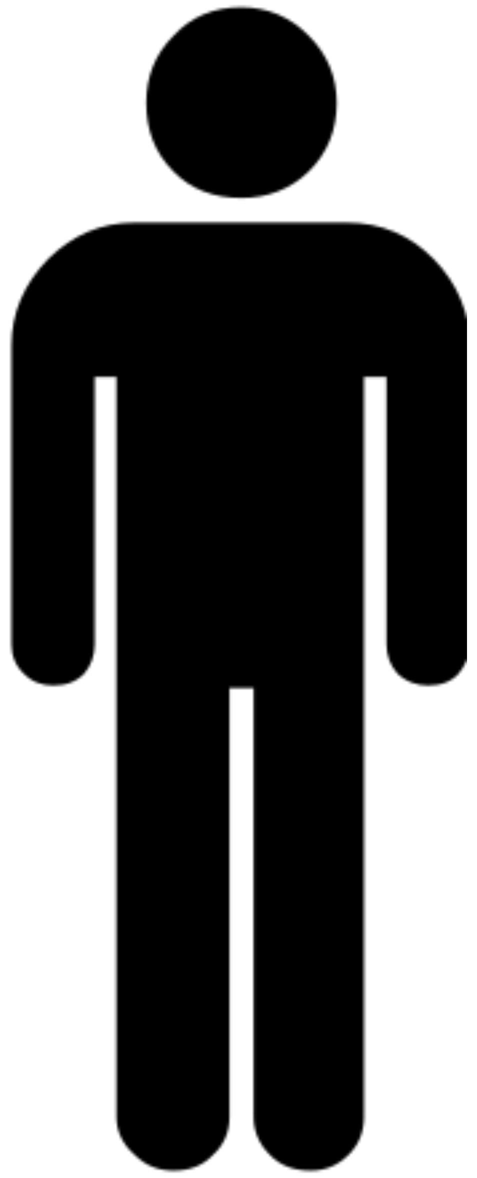
Binary search as a game

Binary search as a game



Binary search as a game

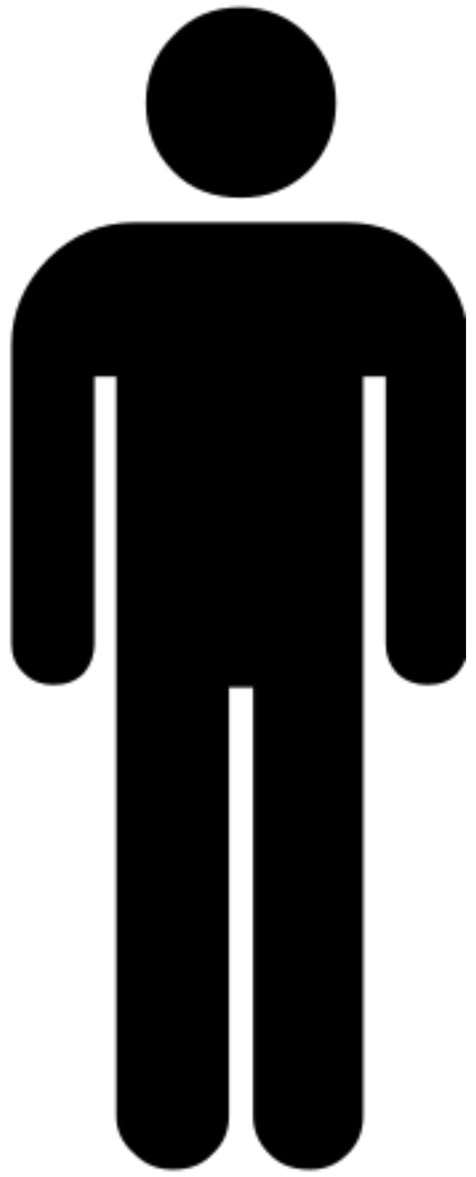
Adversary



No!

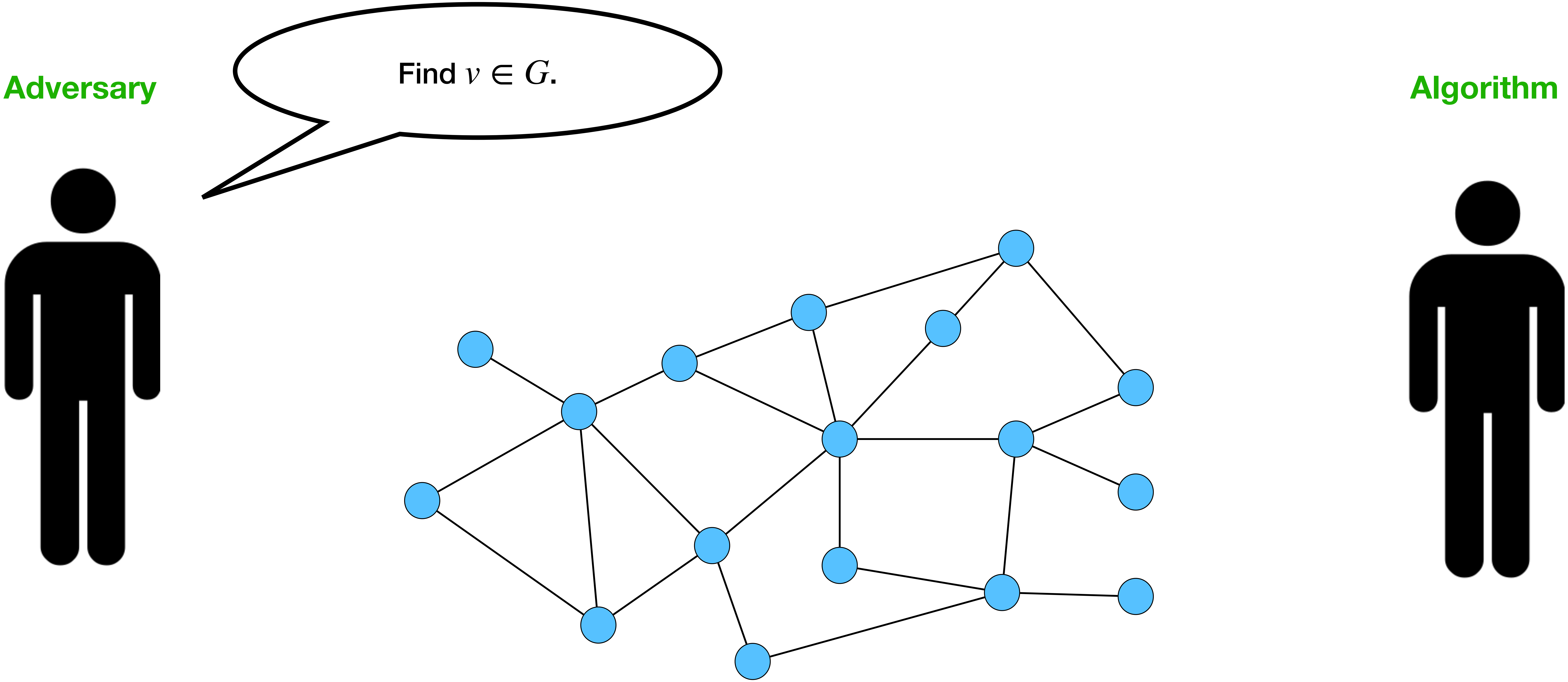
Is $x \leq 250000$?

Algorithm

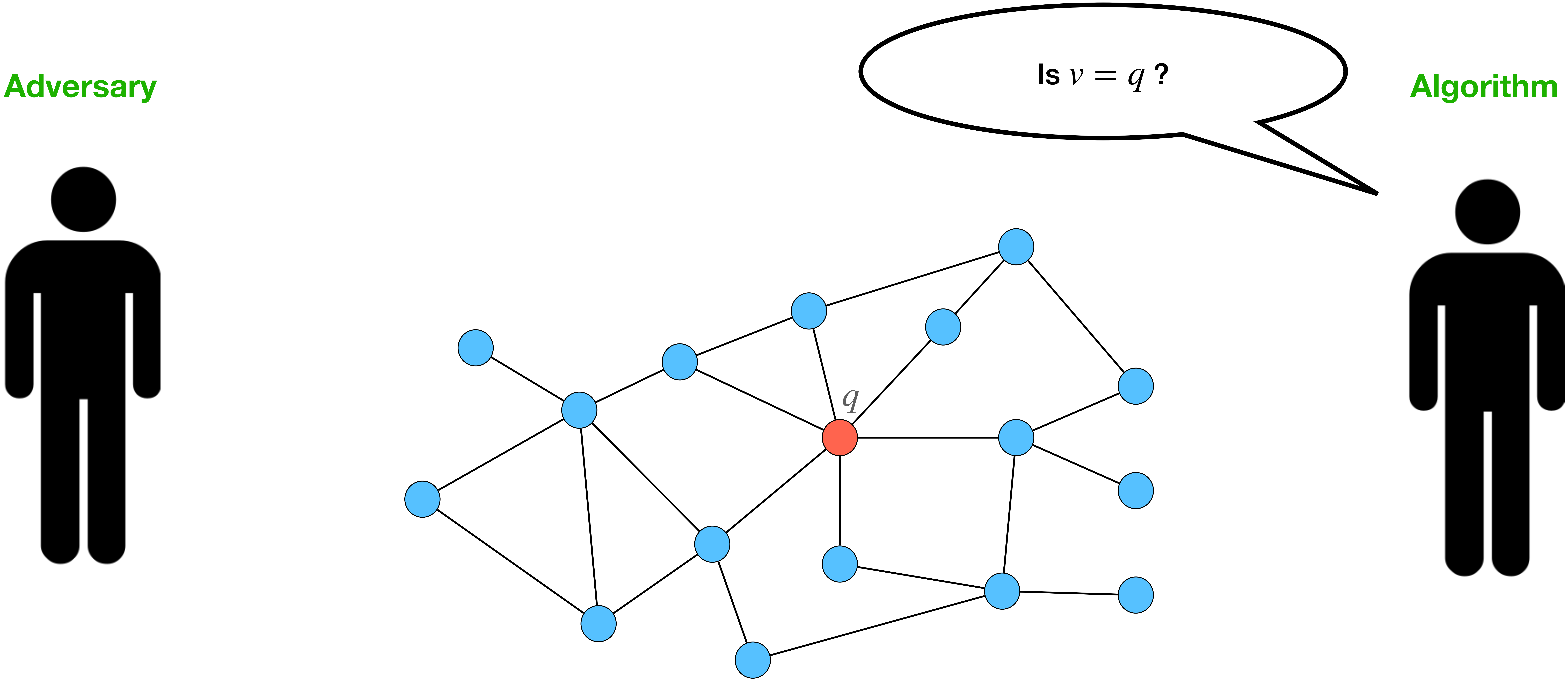


And so on, and so forth…

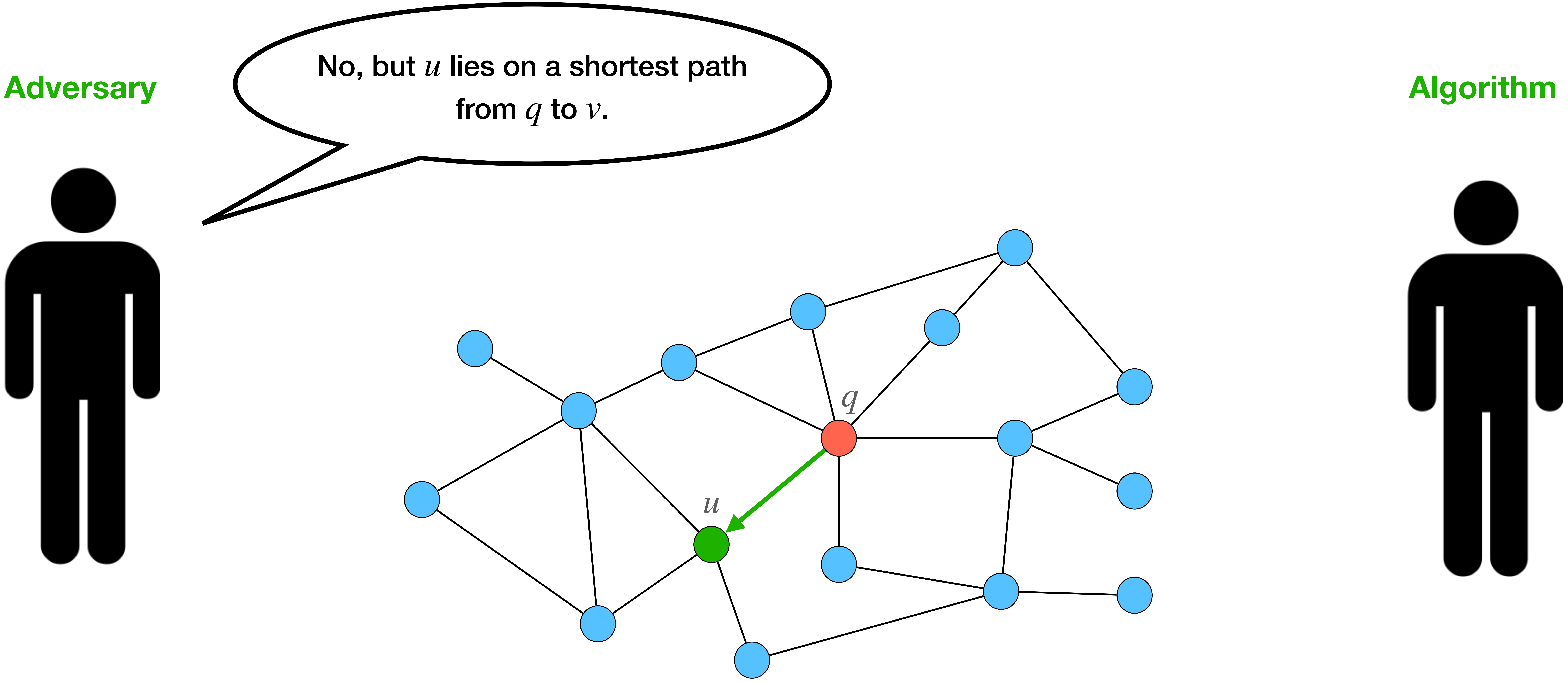
Game on a graph



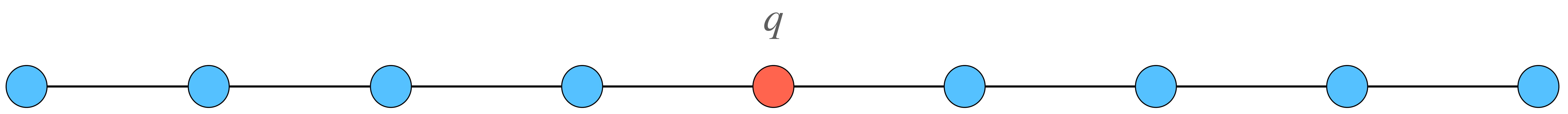
Game on a graph



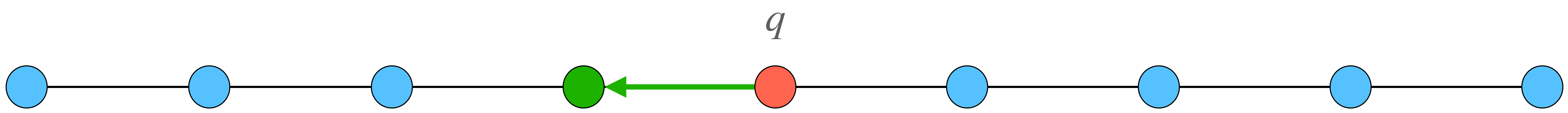
Game on a graph



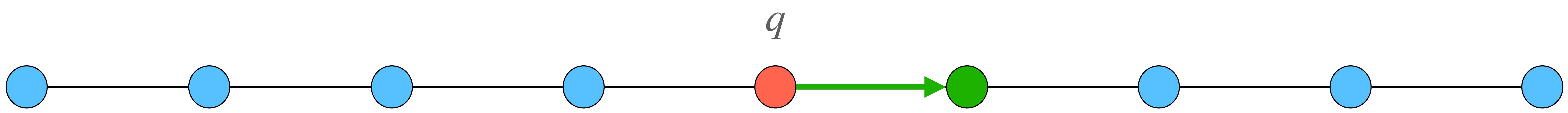
(Almost) a generalization of binary search



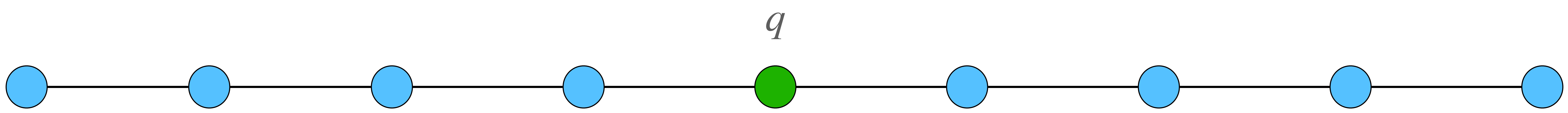
(Almost) a generalization of binary search



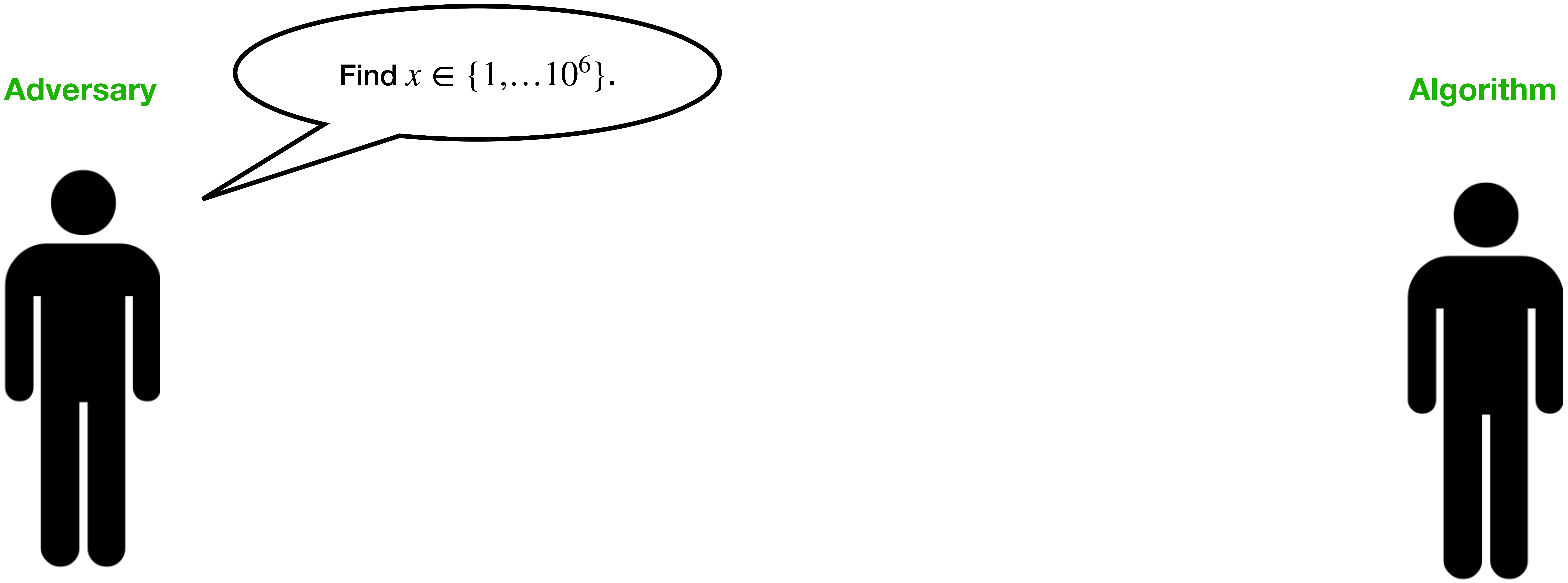
(Almost) a generalization of binary search



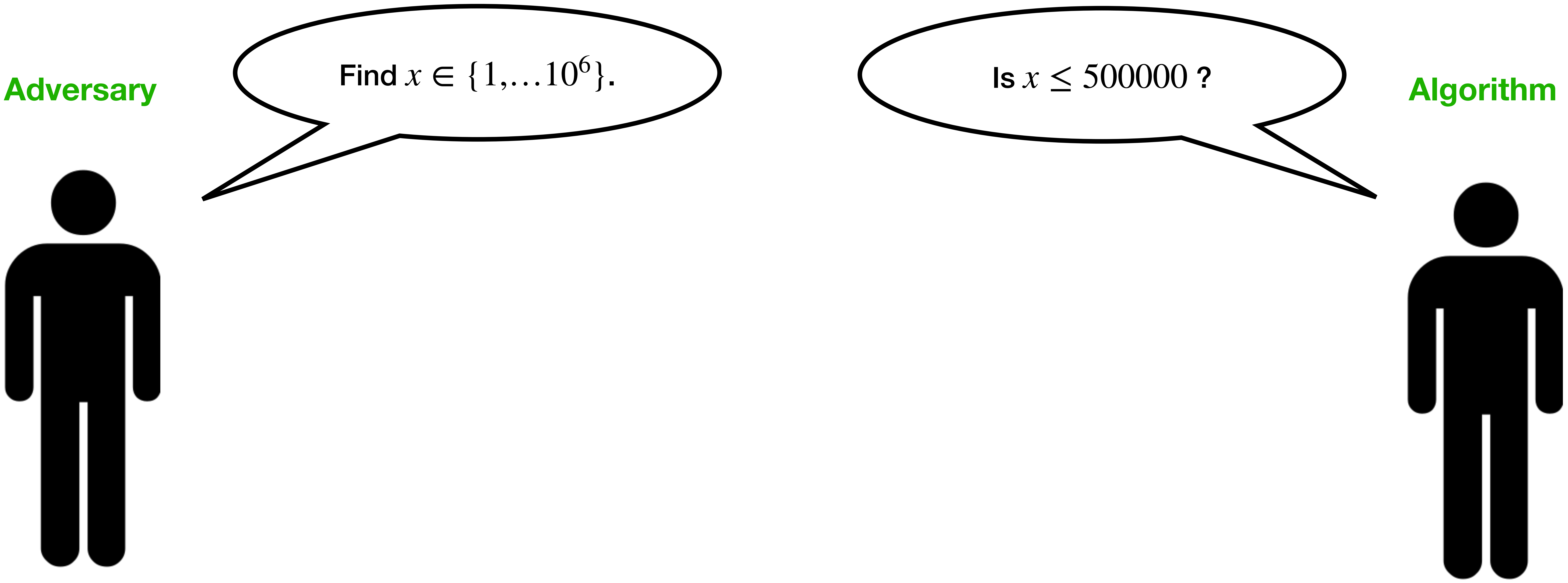
(Almost) a generalization of binary search



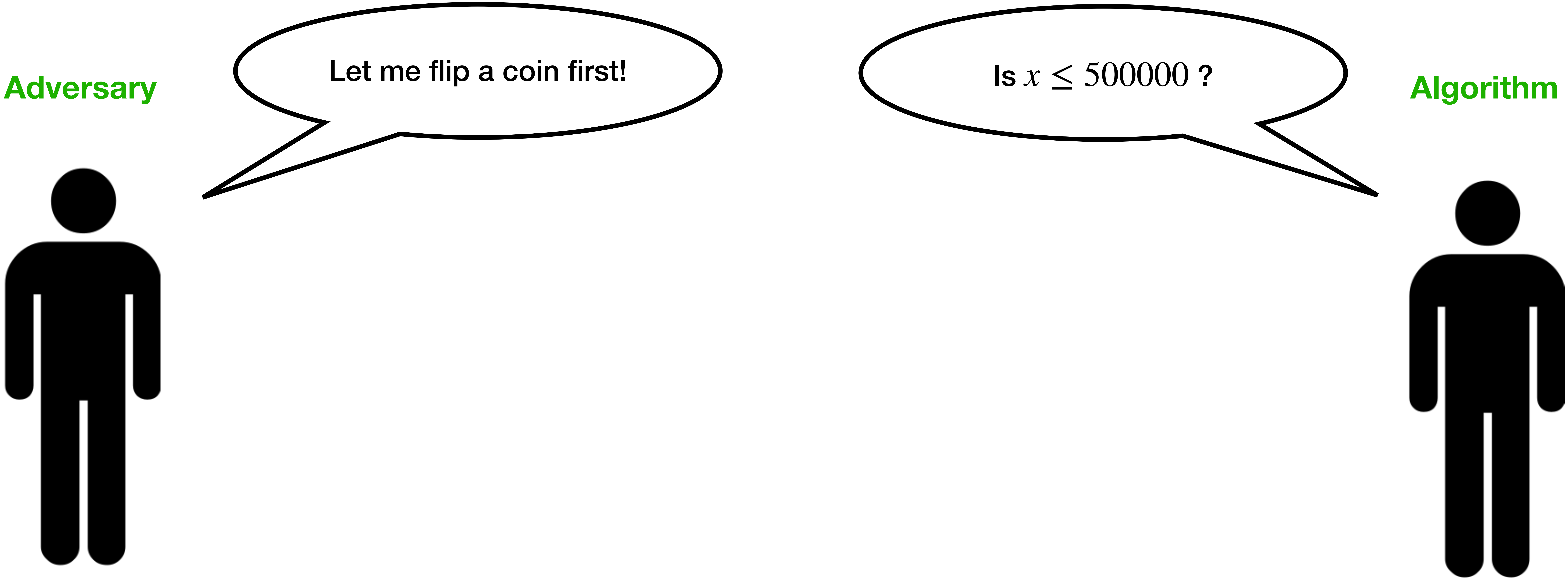
Searching with stochastic noise



Searching with stochastic noise

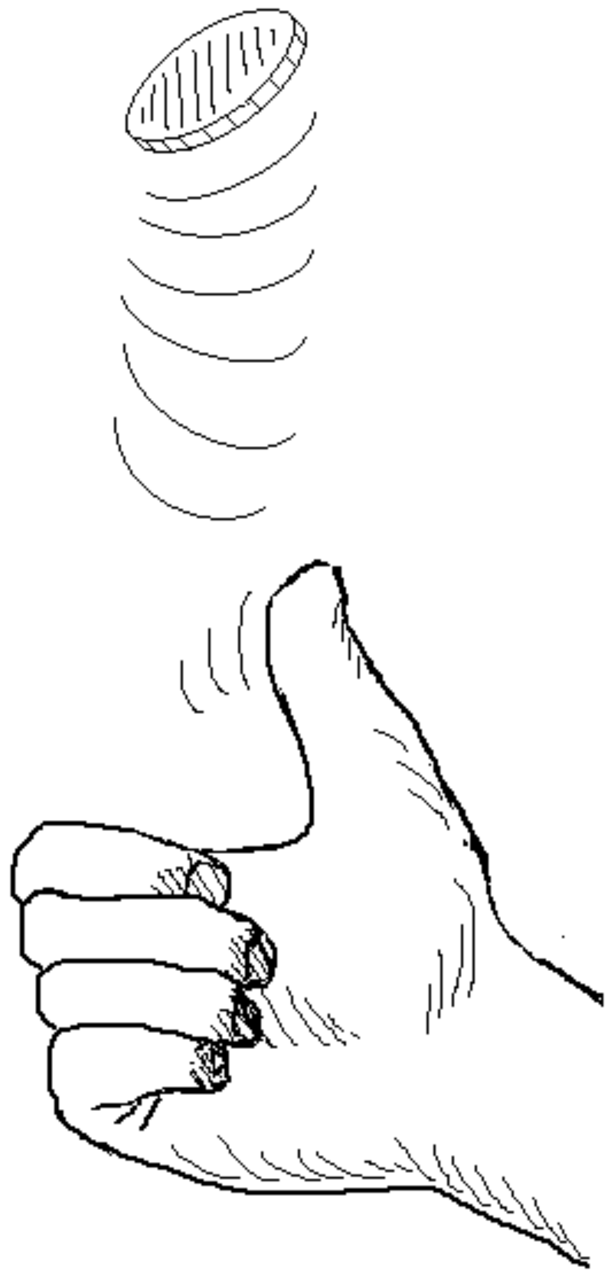


Searching with stochastic noise

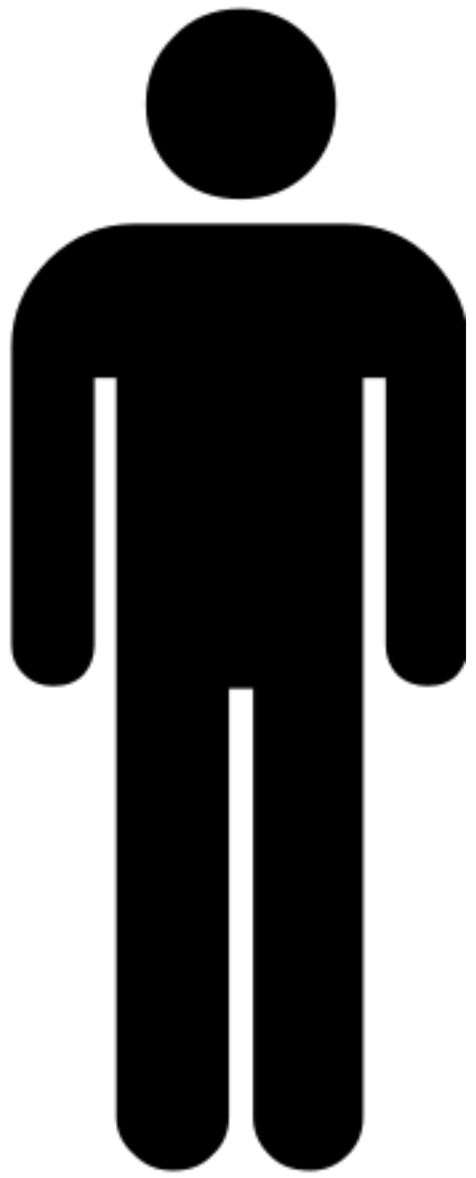


Searching with stochastic noise

Adversary



Algorithm

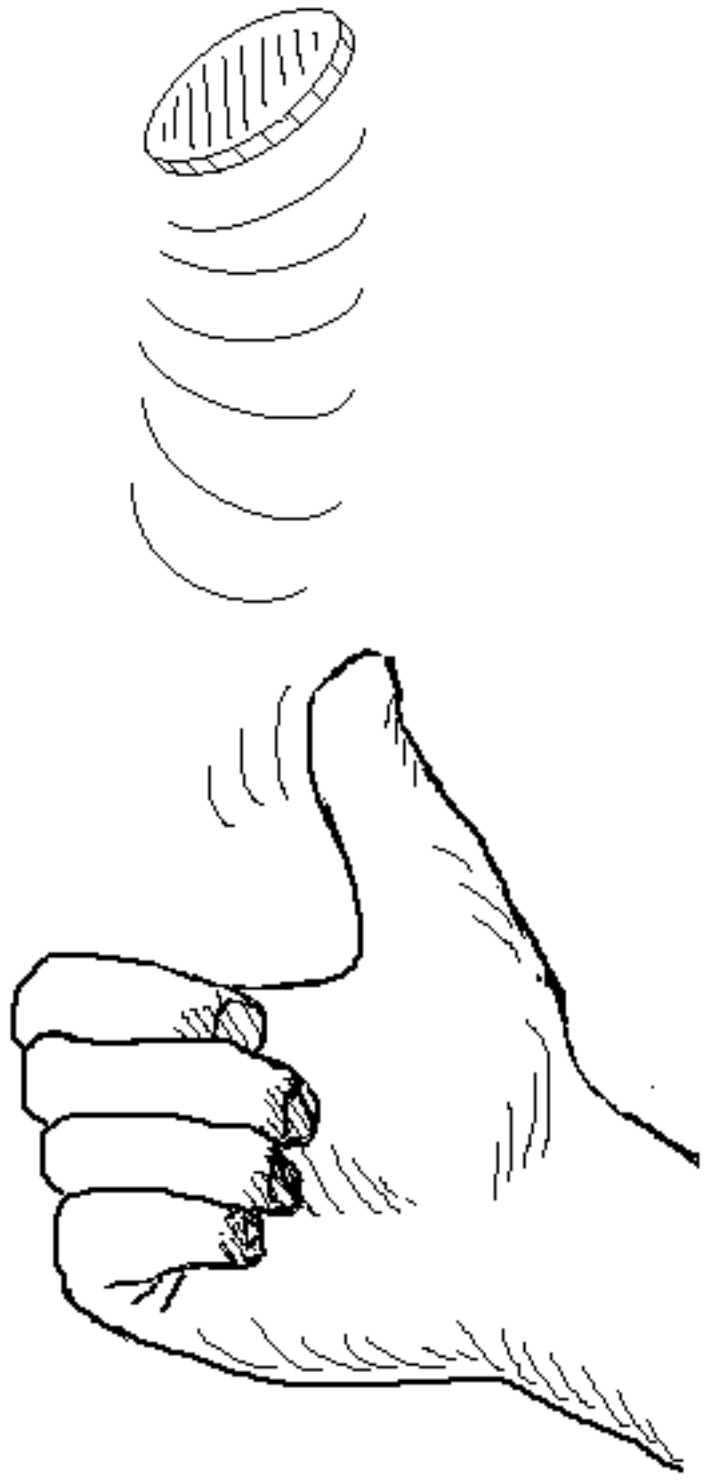


Searching with stochastic noise

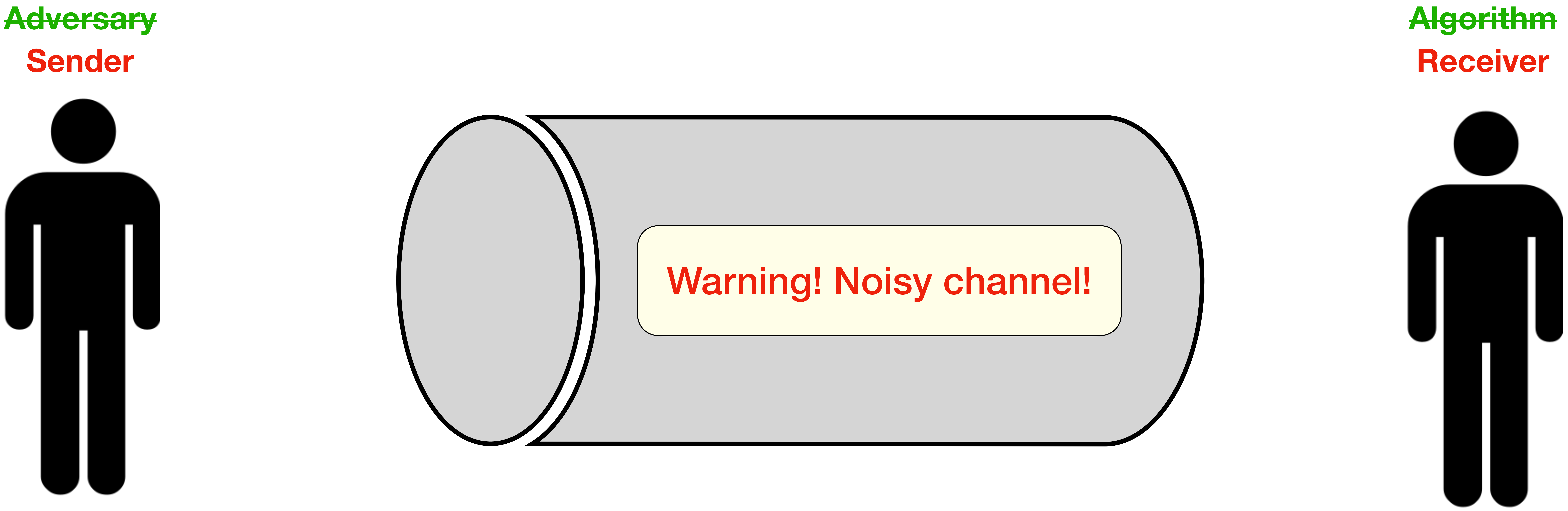
Searching with stochastic noise

Probabilistic errors

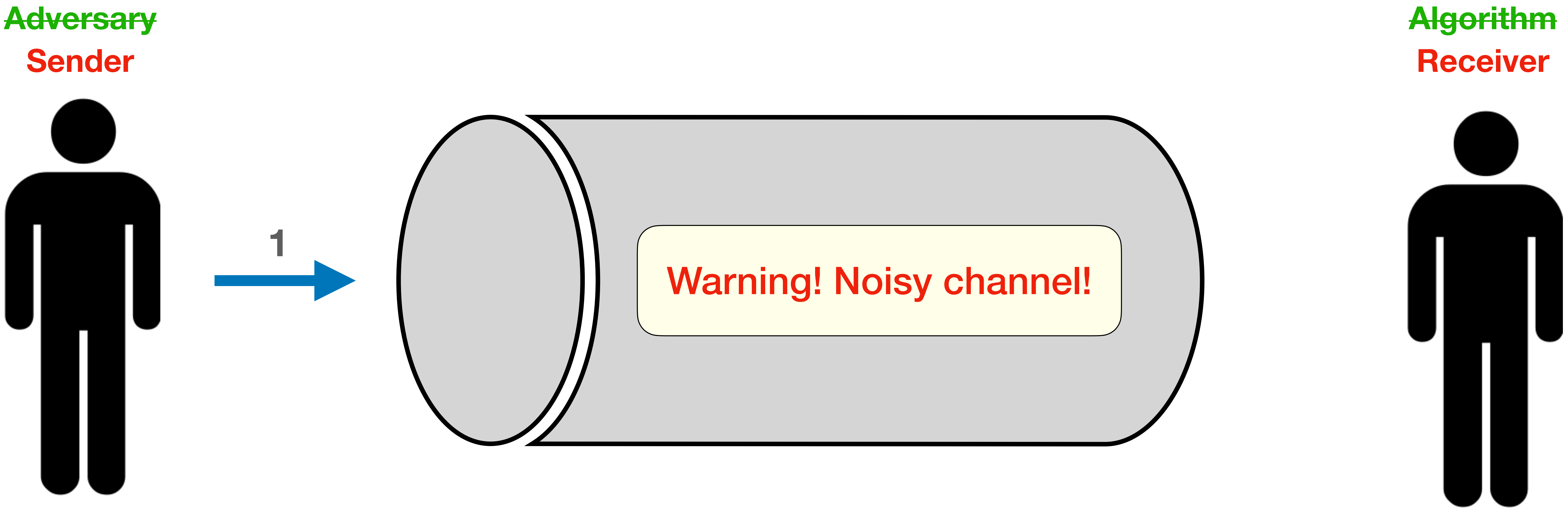
- For each query an erroneous reply independently with probability p, $0 \le p < \frac{1}{2}$.
- We want correct output with probability at least $1 - \delta$.
- We want to know the exact query complexity (we care about constants).



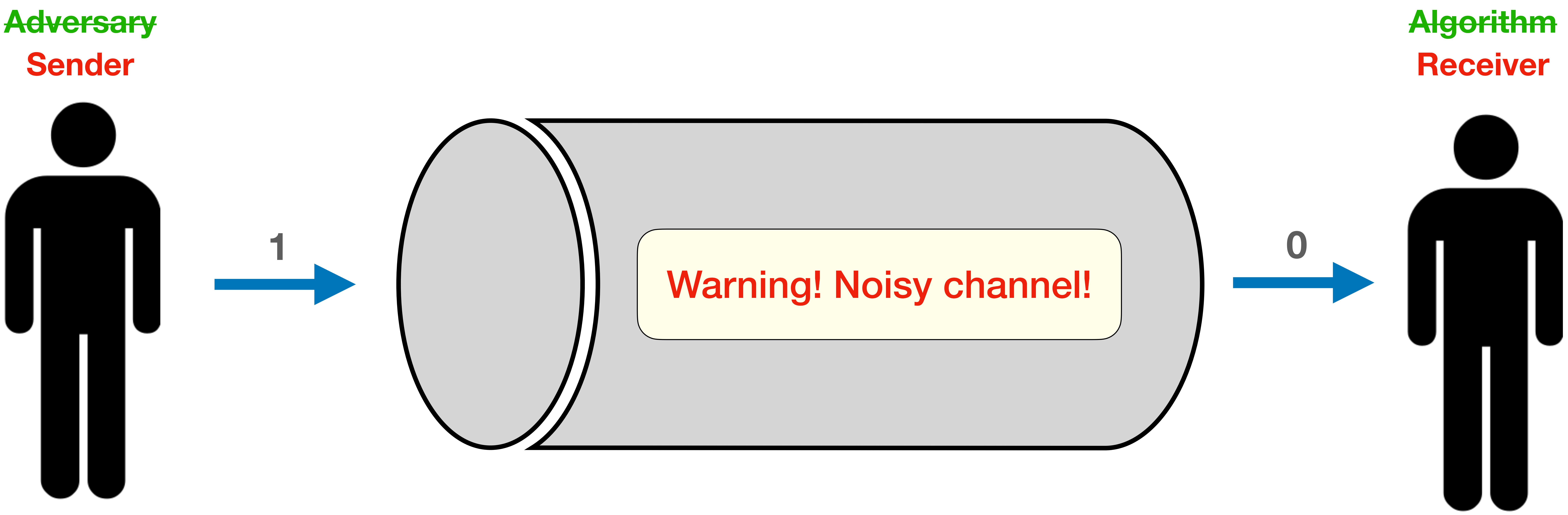
The tale of information theory



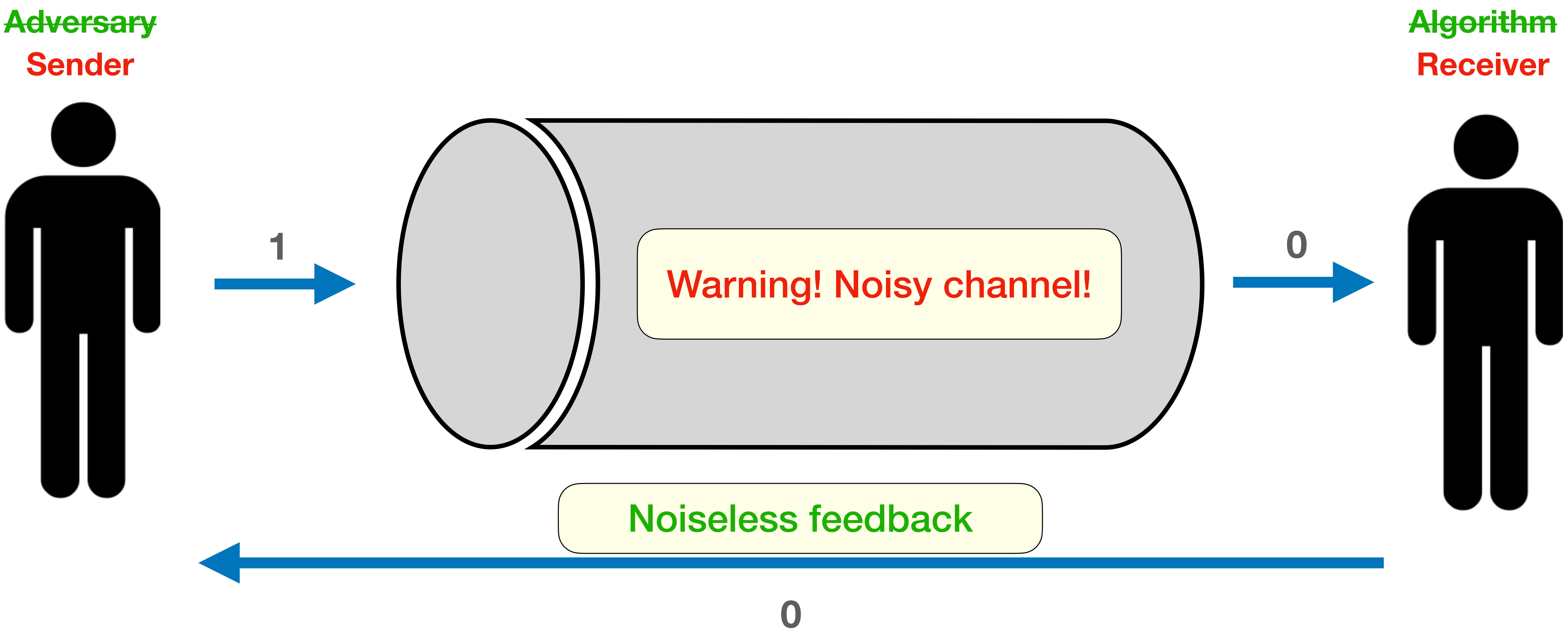
The tale of information theory



The tale of information theory



The tale of information theory



Can't beat roughly

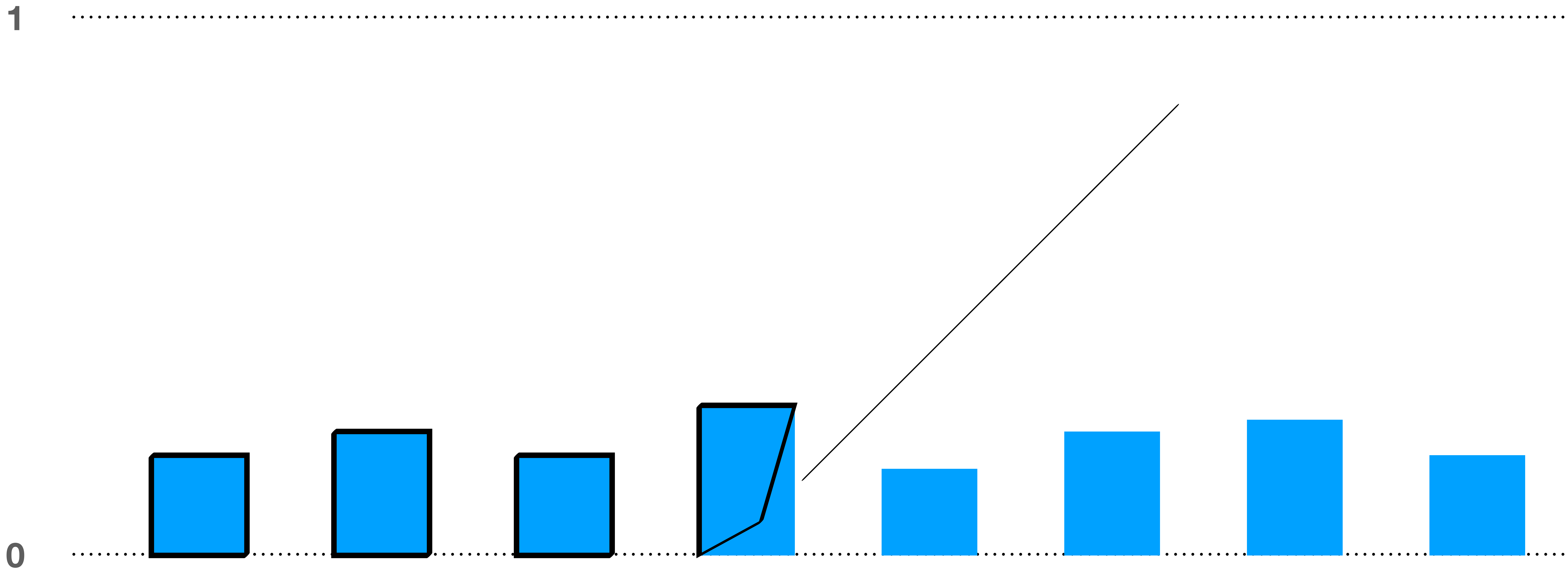
$$\frac{\log_2 n}{1 - H(p)}$$

in expectation.

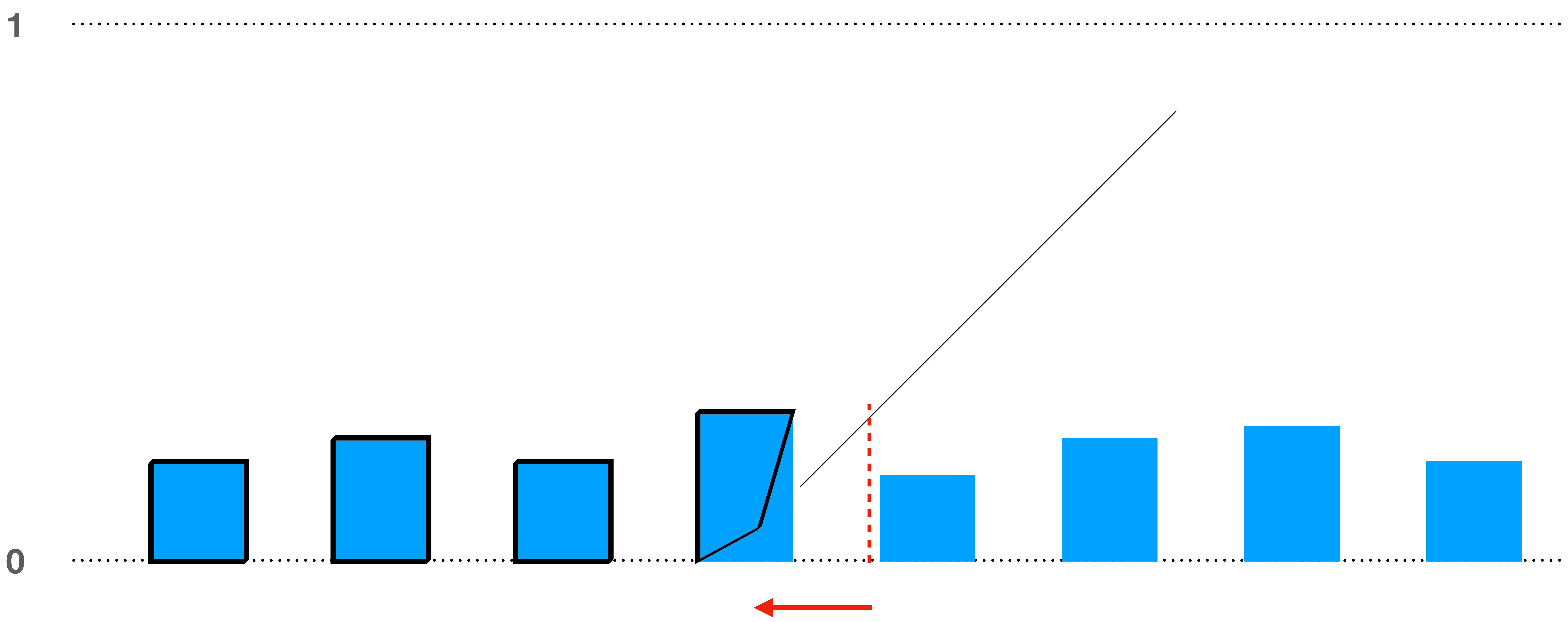
Part II

Intuition and challenges along the way

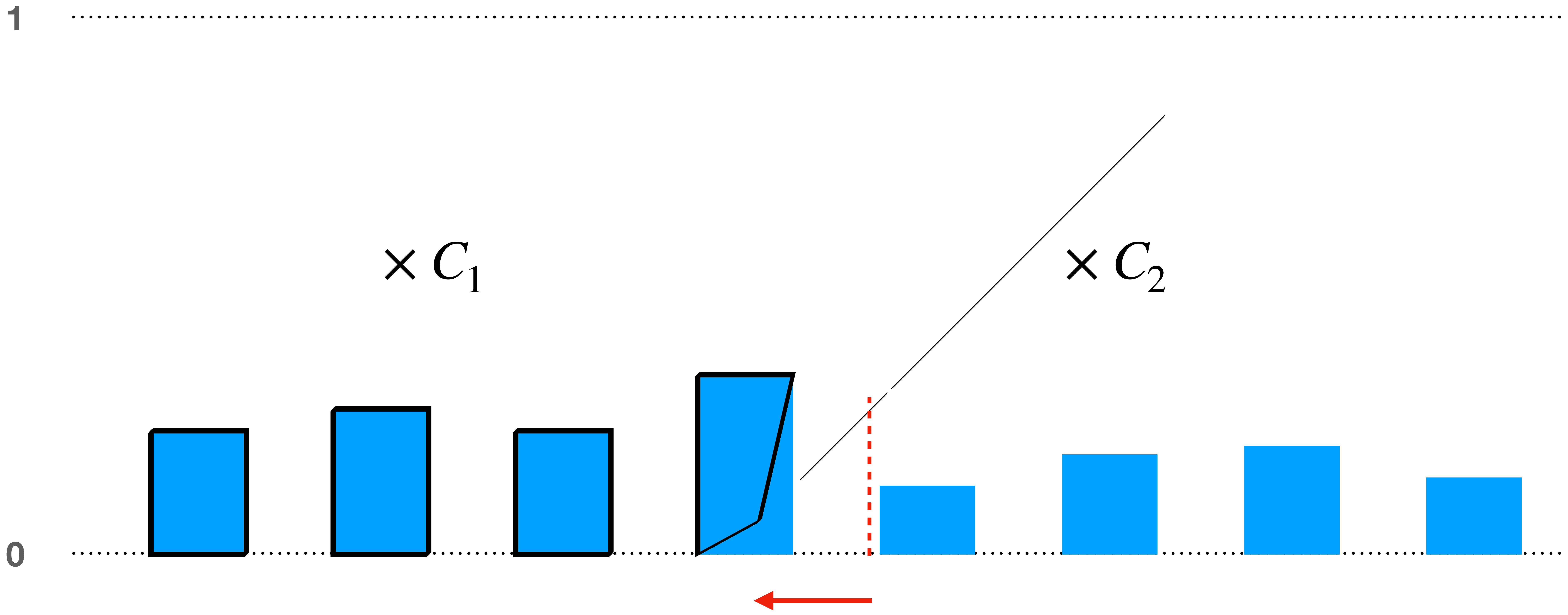
General framework - MWU



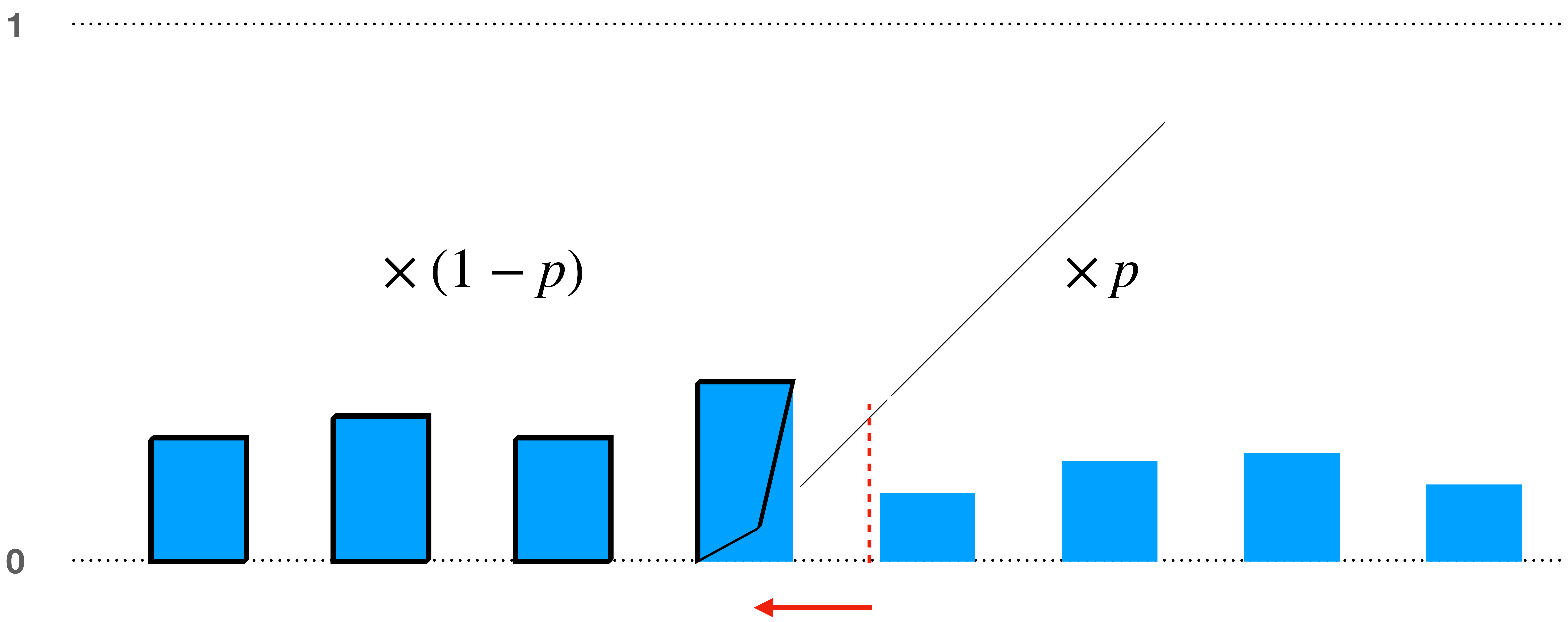
General framework - MWU



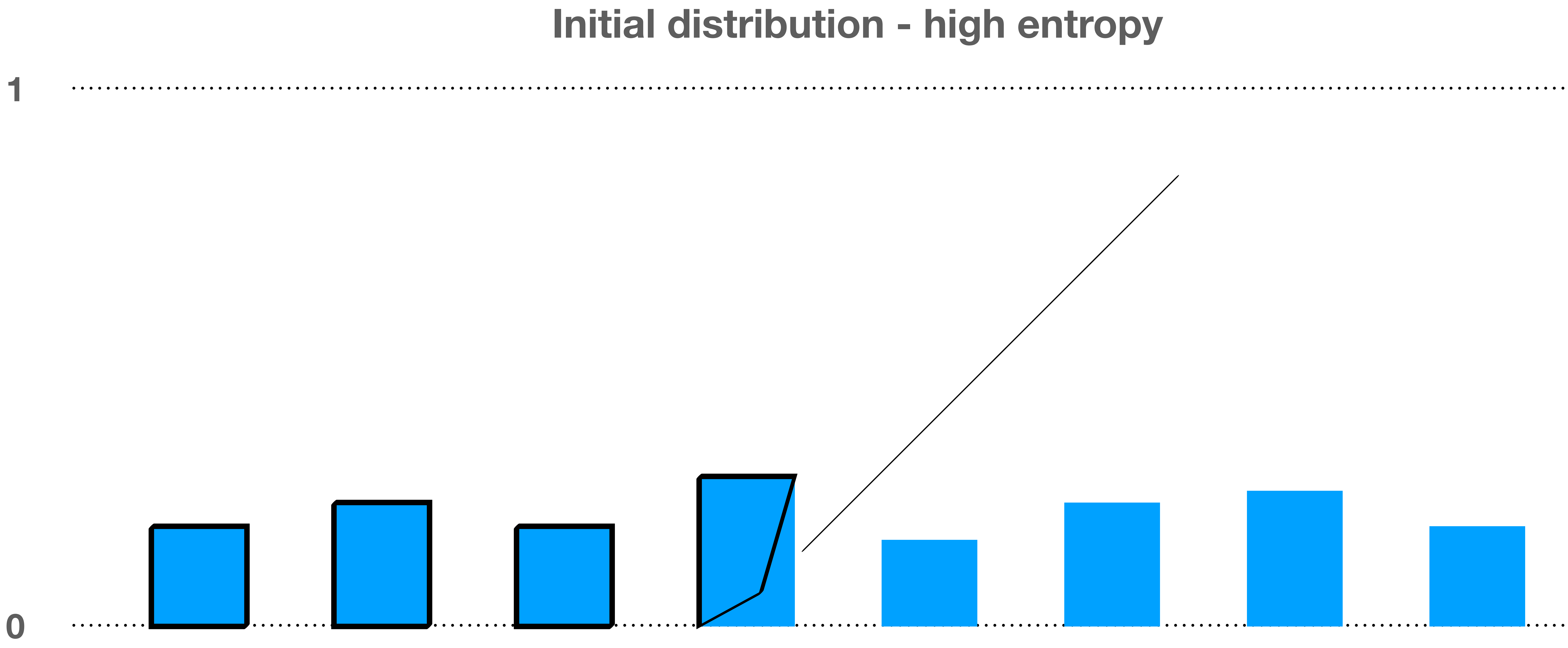
General framework - MWU



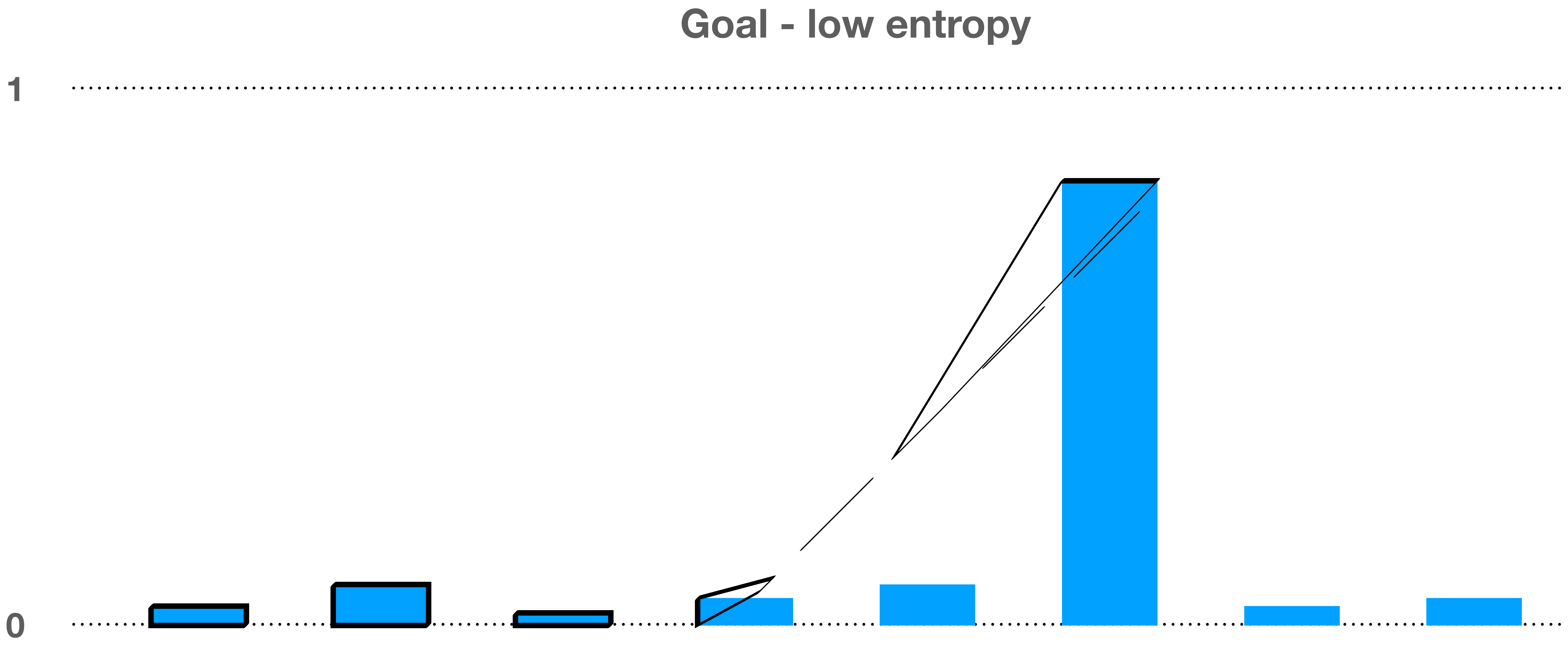
MWU - Bayesian updates



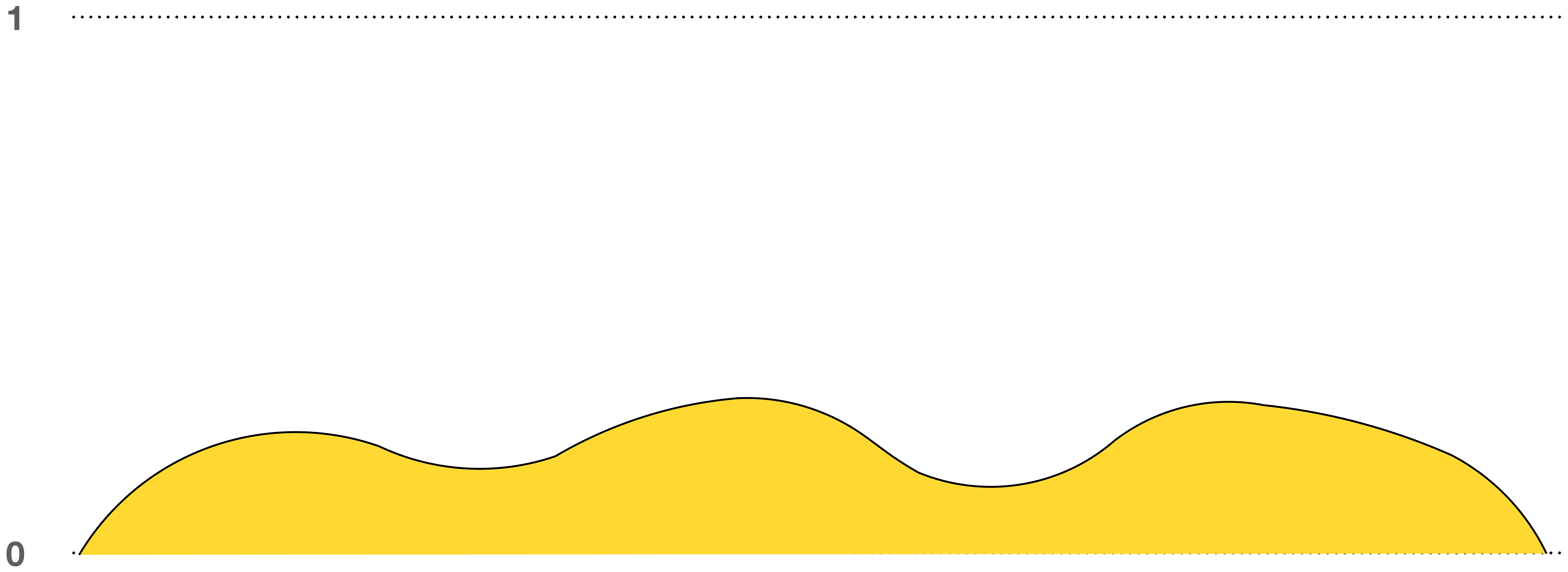
Measuring progress



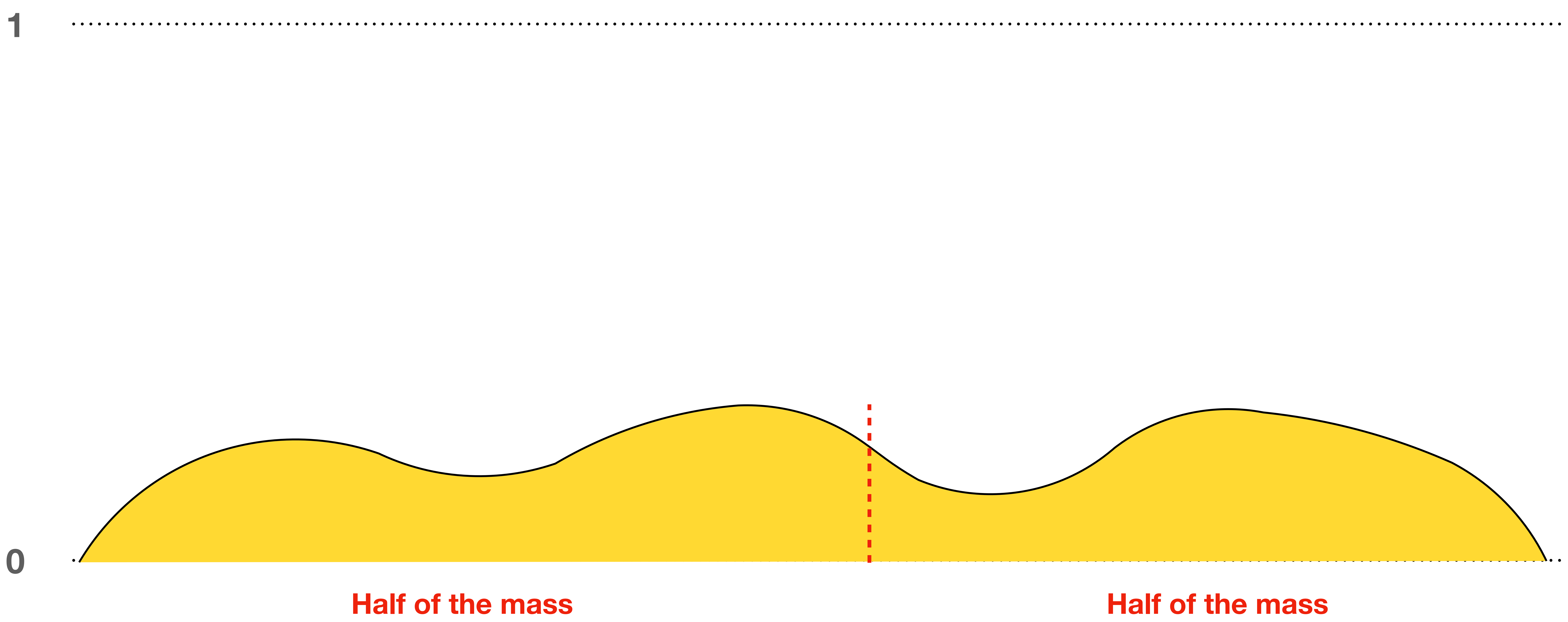
Measuring progress



Idealised scenario [Ben-Or, Hassidim, 2008]

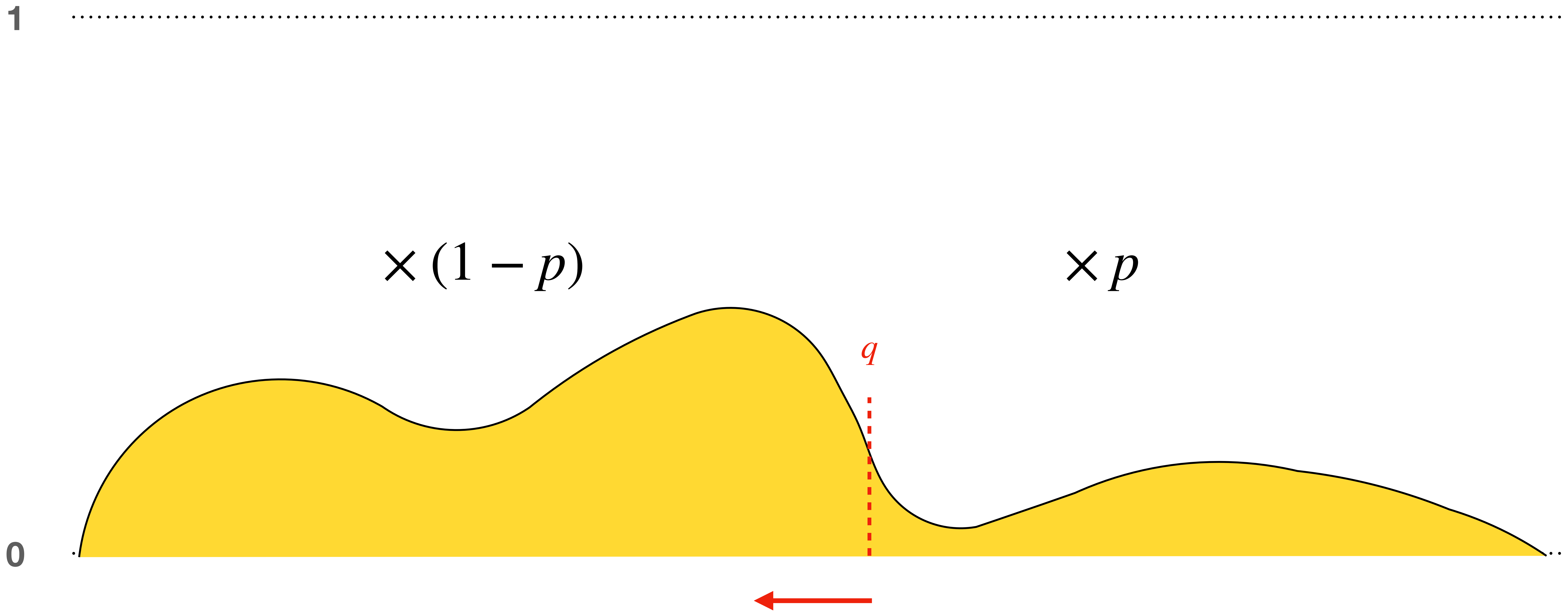


Idealised scenario [Ben-Or, Hassidim, 2008]



Idealised scenario [Ben-Or, Hassidim, 2008]

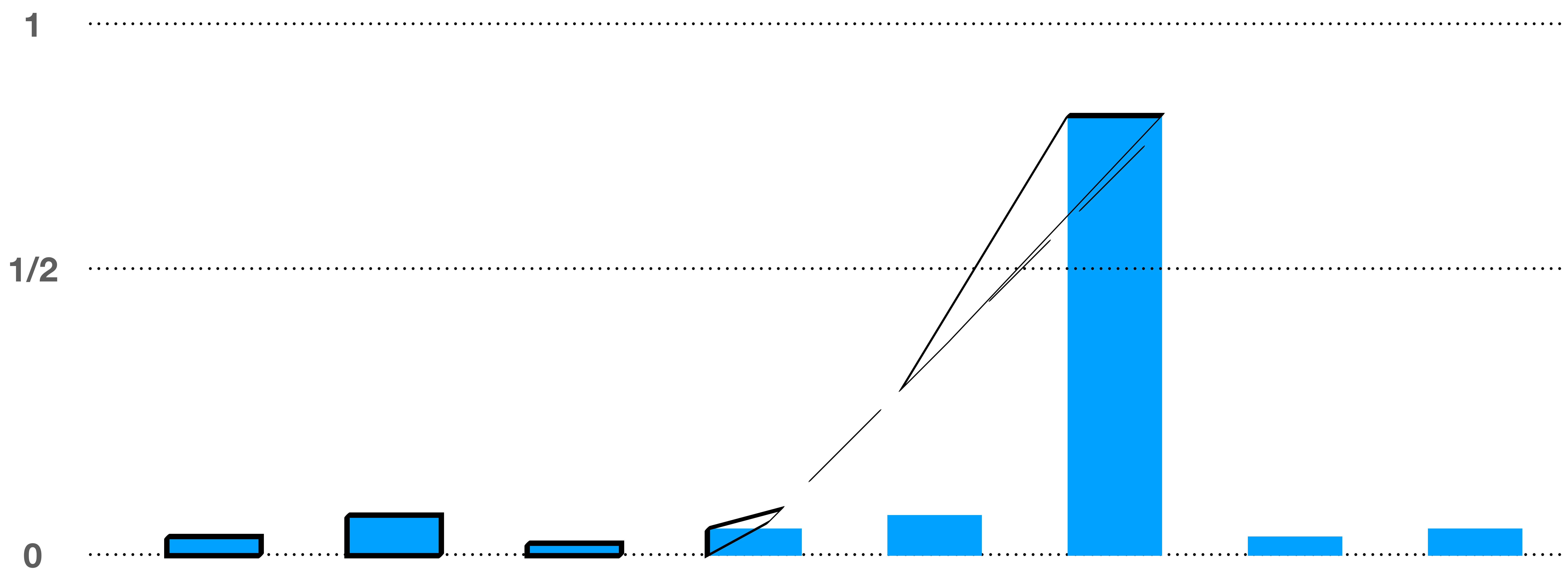
The expected entropy drop is $1 - H(p)$



But in CS we are living in a discrete world!

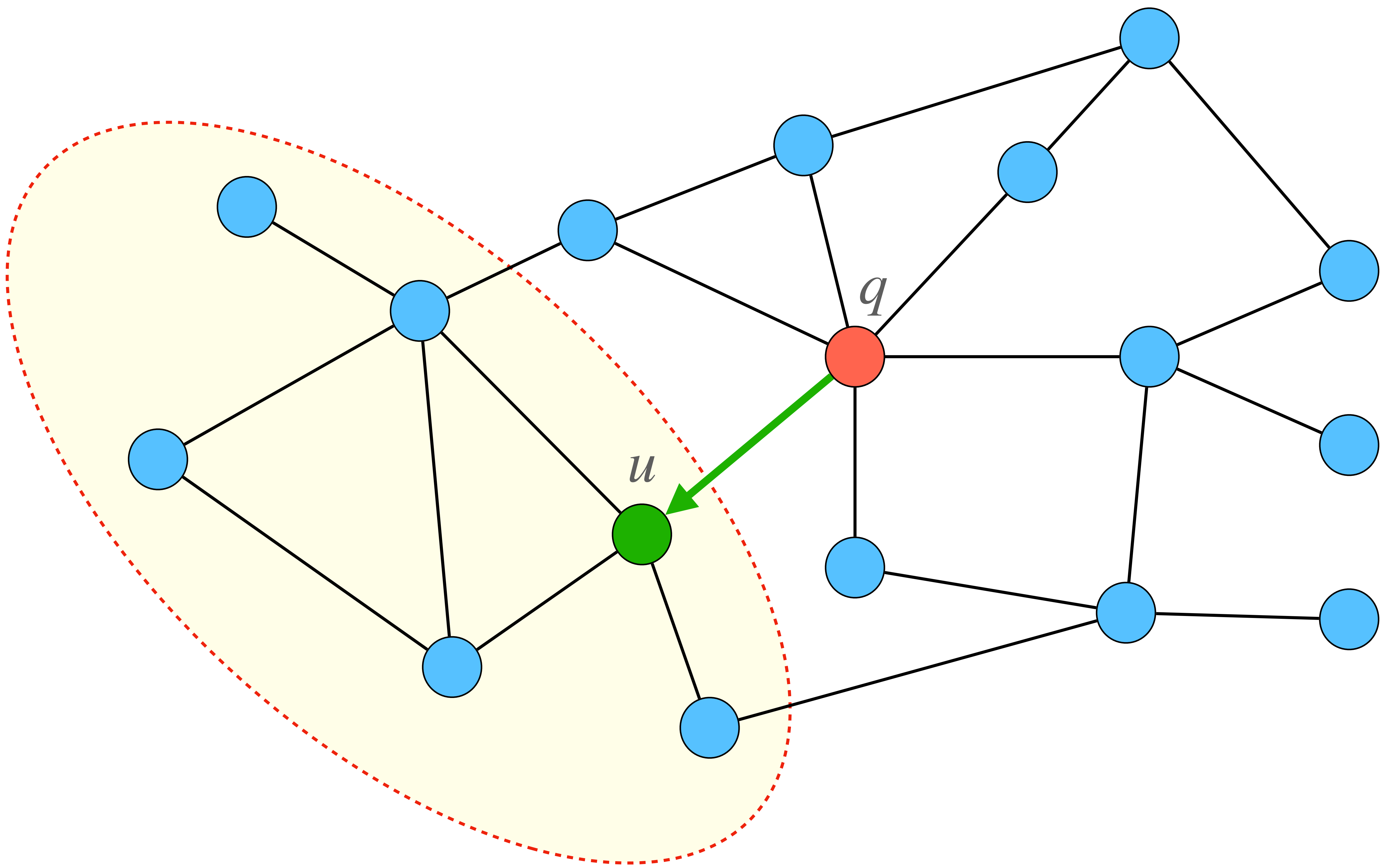
We won’t be able to get a perfect bisection all the time!

It’s a discrete, discrete, discrete world!



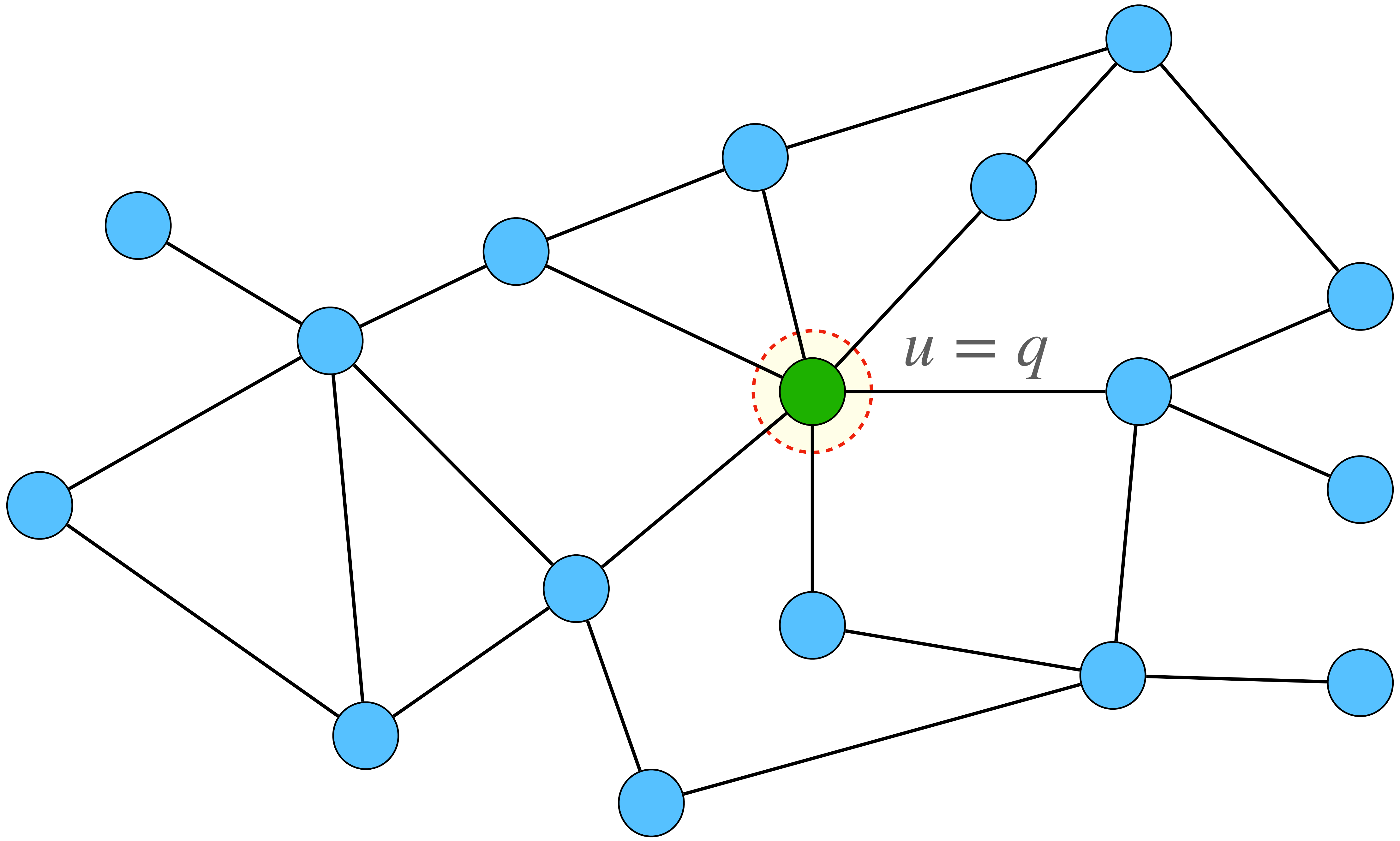
Bisection in graph searching

$N(q, u)$ - set of vertices consistent with query q and response u.



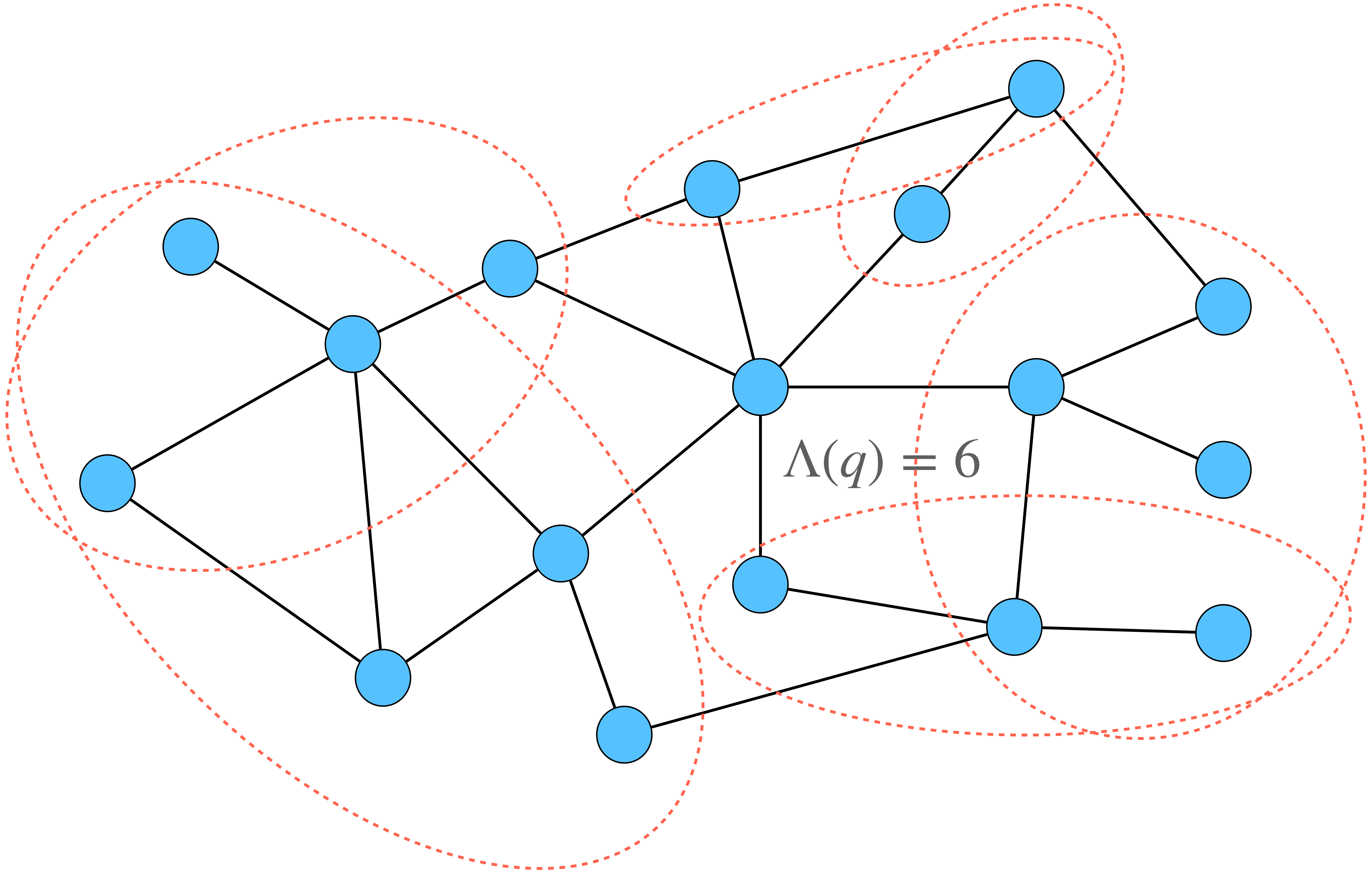
Bisection in graph searching

$N(q, u)$ - set of vertices consistent with query q and response u.



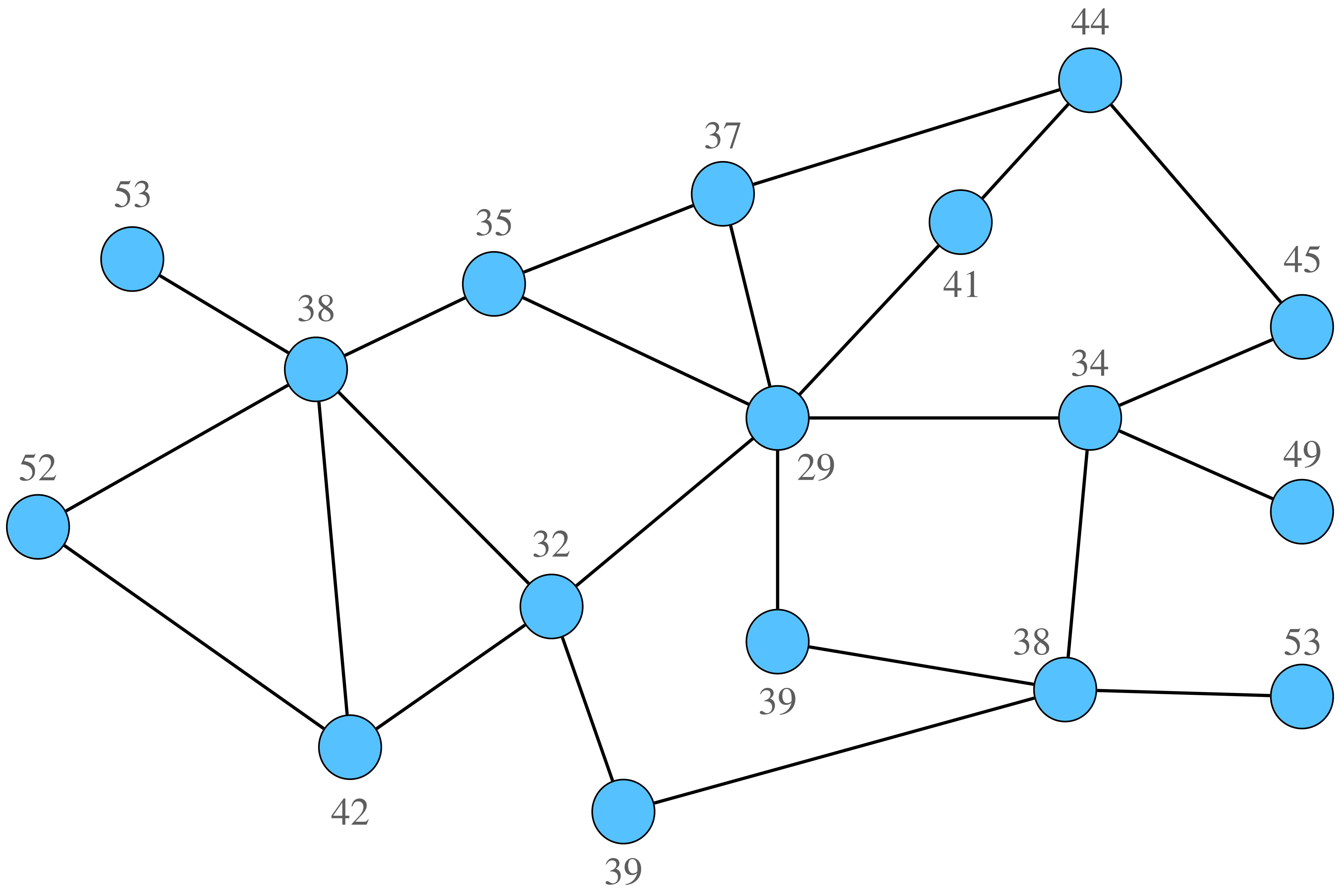
Bisection in graph searching

$$\Lambda(v) = \max_{u \in N(v)} \omega(N(v, u))$$



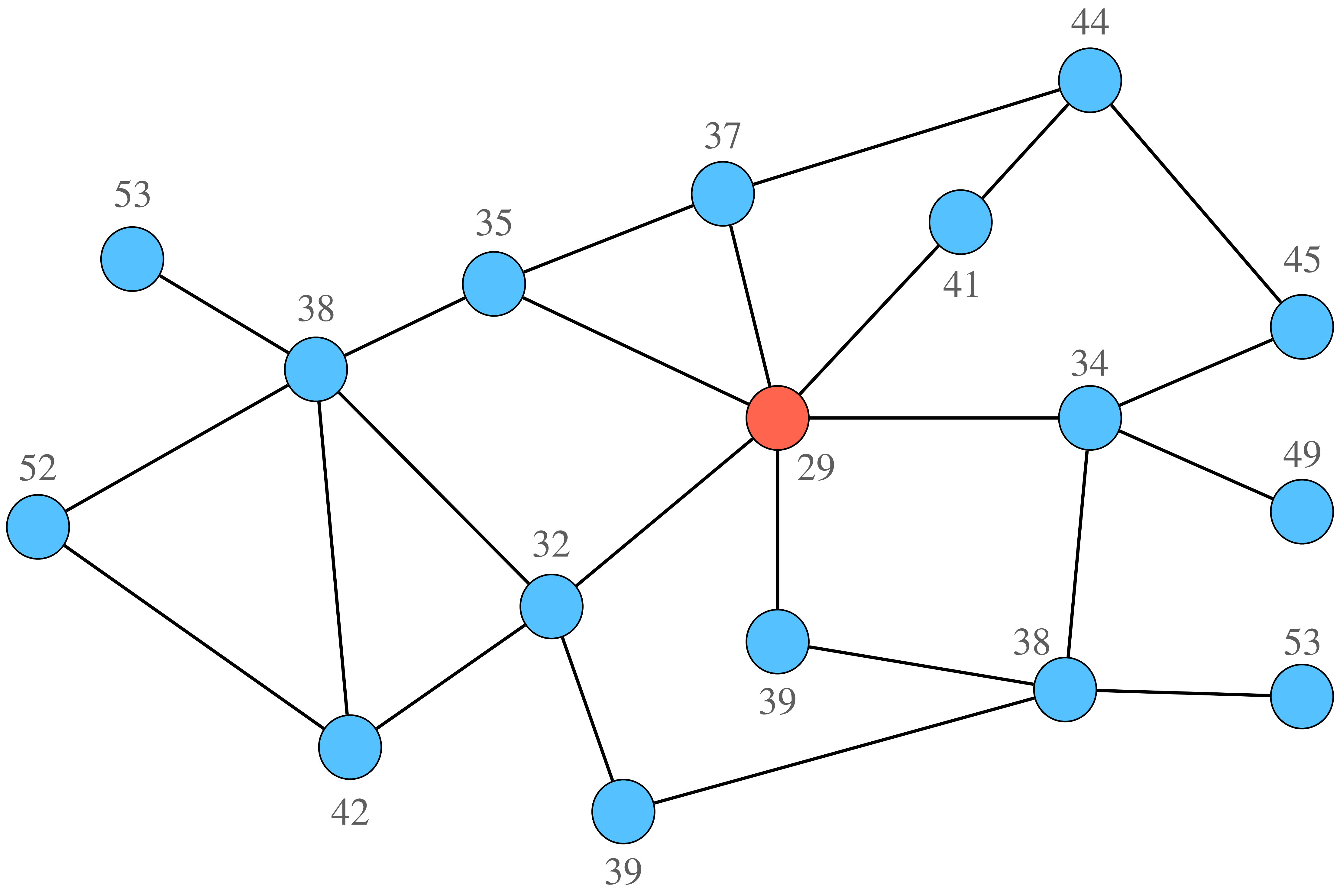
Bisection in graph searching

$$\Phi(v) = \sum_{u \in V} \omega(u) \cdot d(u, v)$$



Bisection in graph searching

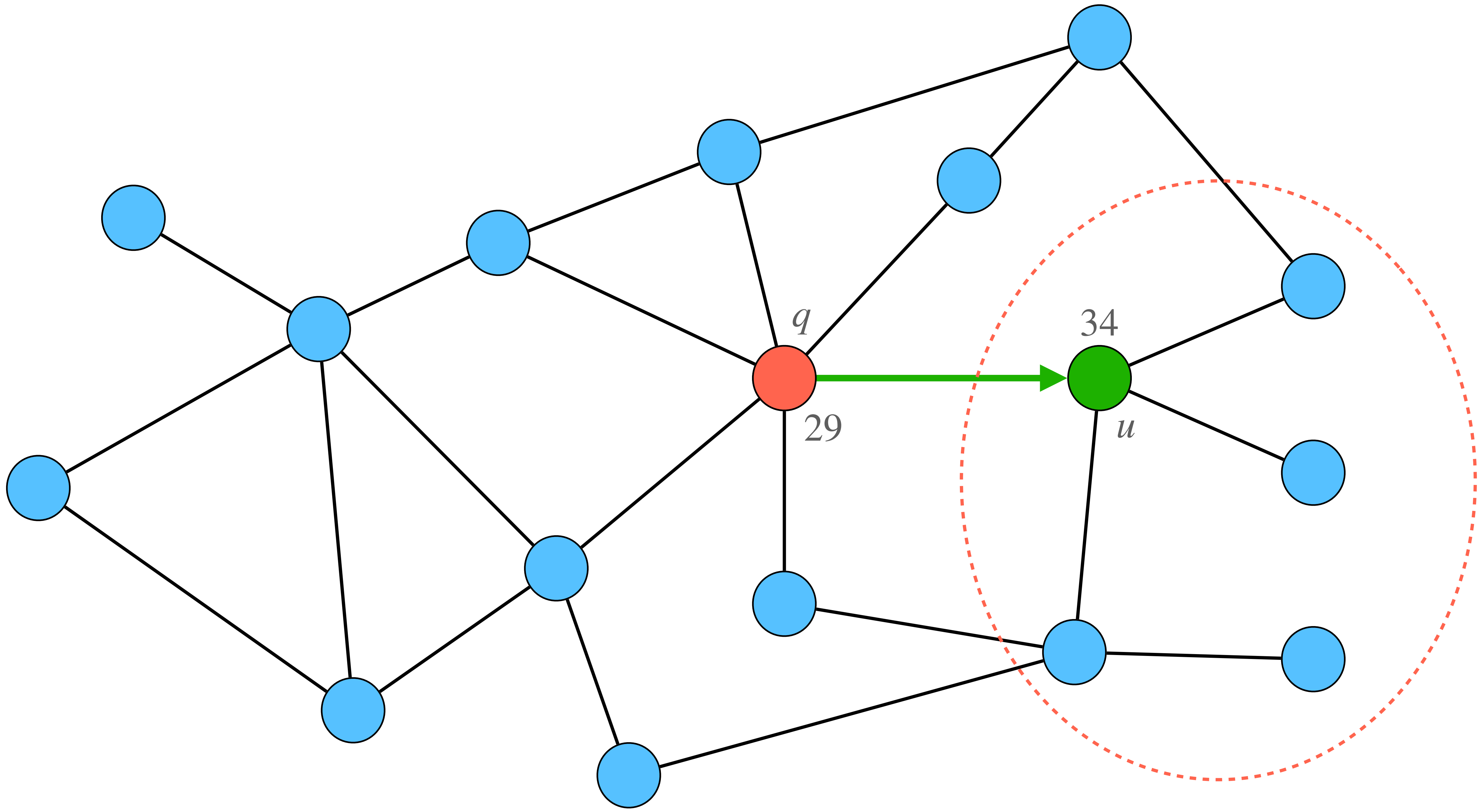
Median of a graph: $q = \arg\min_{v \in V} \Phi(v)$



Bisection in graph searching

Lemma [Emamjomeh-Zadeh et al. '16]:

if q is a median, then $\Lambda(q) \leq \dfrac{\omega(V)}{2}$



Part III

Our results

Previous (close to optimal) results - noisy binary search

Paper	Setting	Query complexity
Ben-Or, Hassidim [FOCS 2008]	Expected	$\frac{1}{1-H(p)}(\log_2 n + O(\log\log n) + O(\log \delta^{-1}))$
Gan et al. [SWAT 2022]	Worst-case Constant p	$\frac{1}{1-H(p)}(\log_2 n + O(\sqrt{\log n \log \delta^{-1}} \log \frac{\log n}{\log \delta^{-1}}) + O(\log \delta^{-1}))$
Gretta, Price [ICALP 2024]	Worst-case More general problem	$\frac{1}{1-H(p)}(\log_2 n + O(\log^{2/3} n \log^{1/3} \delta^{-1}) + O(\log \delta^{-1}))$

Previous (close to optimal) results - noisy binary search

Paper	Setting	Query complexity
Burnashev, Zigangirov [Problemy Peredachi Informatsii, 1974]	Expected Target unif. random	$\frac{1}{1-H(p)}(\log_2 n + \log_2 \delta^{-1} + \log_2 \frac{1-p}{p})$
Ben-Or, Hassidim [FOCS 2008]	Expected	$\frac{1}{1-H(p)}(\log_2 n + O(\log\log n) + O(\log \delta^{-1}))$
Gan et al. [SWAT 2022]	Worst-case Constant p	$\frac{1}{1-H(p)}(\log_2 n + O(\sqrt{\log n \log \delta^{-1}} \log \frac{\log n}{\log \delta^{-1}}) + O(\log \delta^{-1}))$
Gretta, Price [ICALP 2024]	Worst-case More general problem	$\frac{1}{1-H(p)}(\log_2 n + O(\log^{2/3} n \log^{1/3} \delta^{-1}) + O(\log \delta^{-1}))$

Previous results - noisy graph search

Paper	Setting	Query complexity
Emamjomeh-Zadeh et al. [STOC 2016]	Expected	$\frac{1}{1-H(p)}(\log_2 n + o(\log n) + O(\log^2 \delta^{-1}))$
Dereniowski et al. [SOSA 2019]	Worst-case	$\frac{1}{1-H(p)}(\log_2 n + O(\sqrt{\log n \log \delta^{-1}} \log \frac{\log n}{\log \delta^{-1}}) + O(\log \delta^{-1}))$

Our results

Model	Binary search	Graph search
Worst case complexity	$\frac{1}{1-H(p)}(\log_2 n + O(\sqrt{\log n \log \delta^{-1}}) + O(\log \delta^{-1}))$	
Expected complexity	$\frac{1}{1-H(p)}(\log_2 n + O(\log \delta^{-1}))$	$\frac{1}{1-H(p)}(\log_2 n + O(\log \log n) + O(\log \delta^{-1}))$

Contributions

- **Simple** - query the median until its over.
- **Fast** - close to optimal, slightly improving previous complexities.
- **Correct** - we correct errors from previous literature.

Techniques

- New measure of progress (in analysis): total weight minus the weight of the heaviest element.
- Random choice of a query [Burnashev, Zigangirov, 1974] with tighter analysis.

Open problems

- Closing the small gap between lower and upper bounds.
- Graph (and tree) searching with permanent probabilistic noise, see [Boczkowski et al. ESA 2018].
- Improving lower order terms in noisy binary search with monotonic probabilities, see [Greta & Price, ICALP 2024].