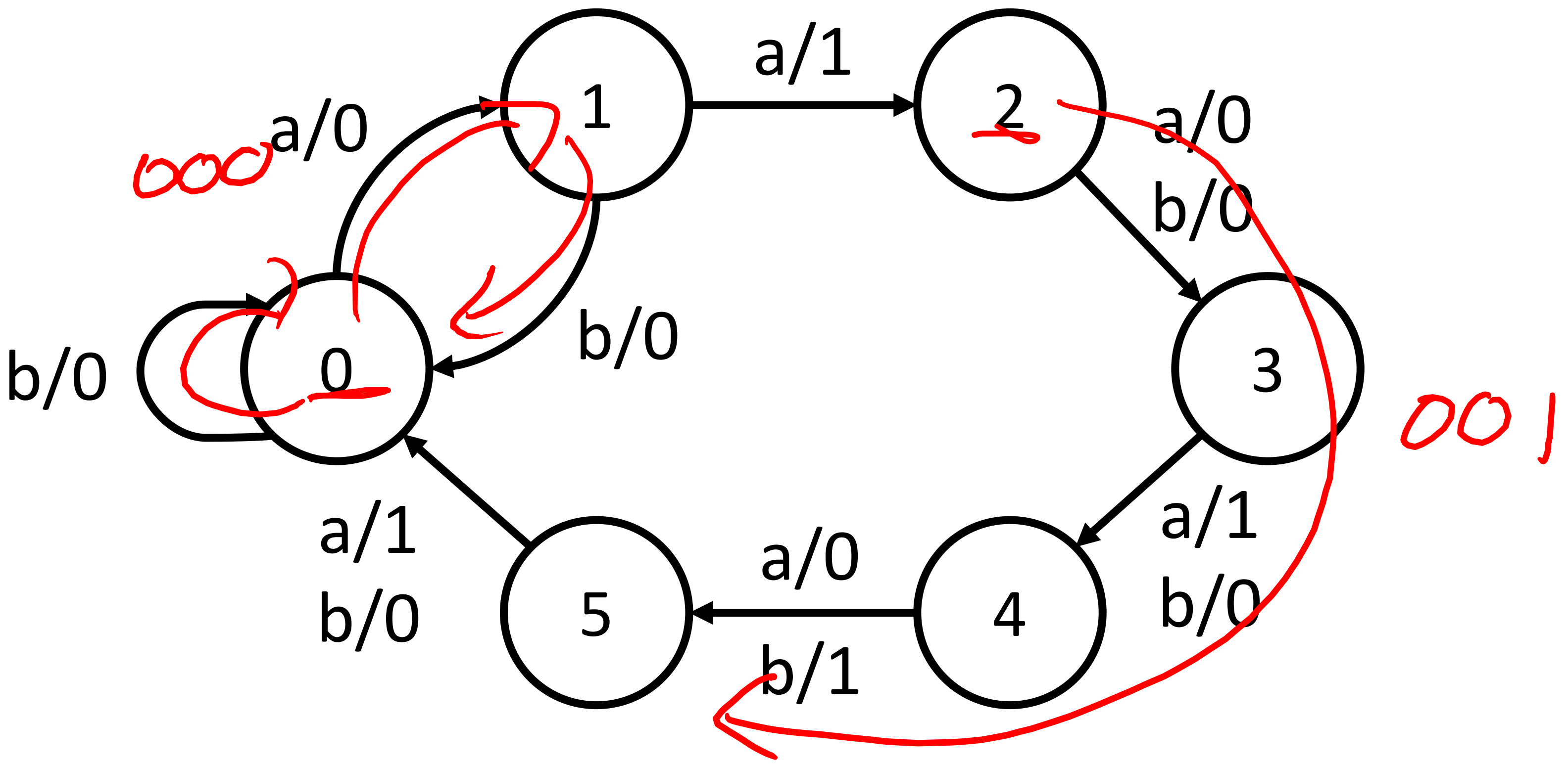
Minimal Separating Sequences for All Pairs of States in $O(m \log n)$

Rick Smetsers **Joshua Moerman** David N. Jansen

Dutch Model Checking Day 2018

Separating Sequences



- Sequences which give different outputs on different states
- Minimal: no shorter separating sequence exists
- Motivation: black-box conformance testing

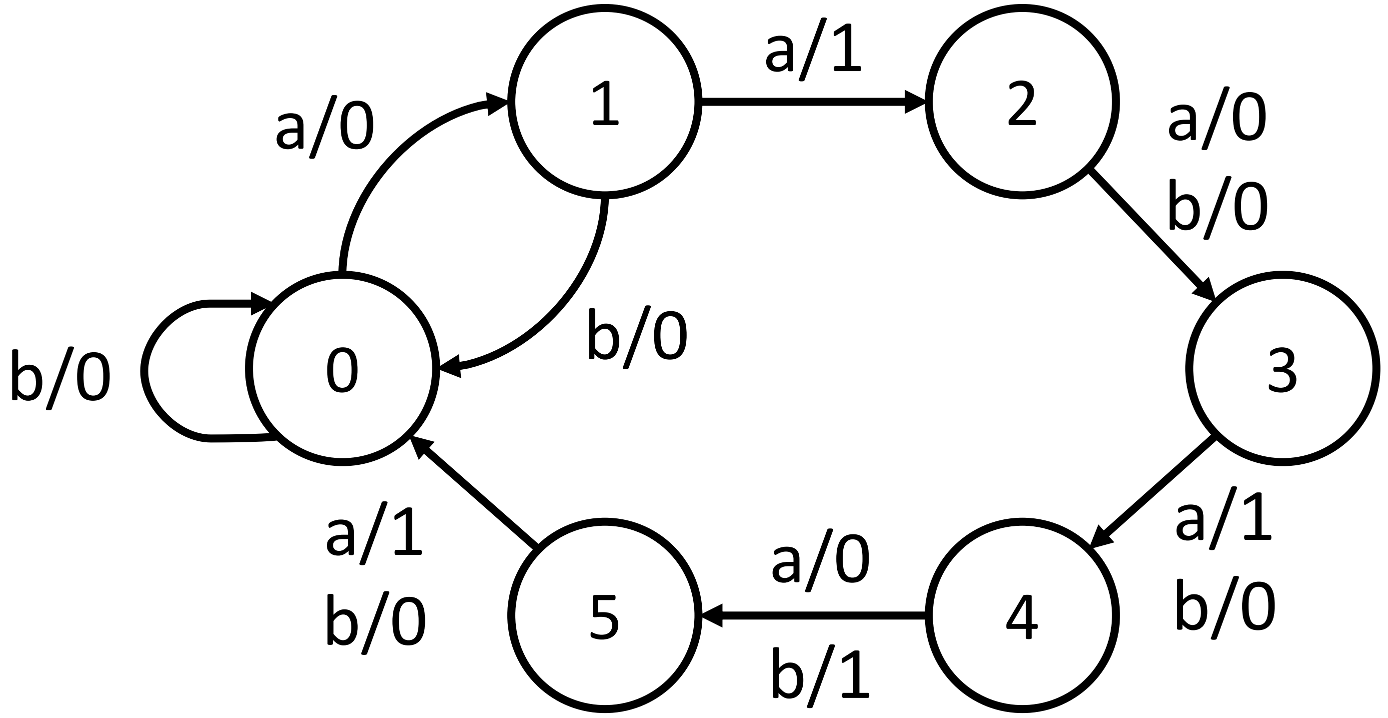
… for All Pairs of States!

- Classic problem in automata theory
- Doing bisimulation for all pairs would require $O(mn^2\alpha(n))$
- Partition refinement gives $O(mn)$
- We extend Hopcroft's $O(m \log n)$ algorithm to return minimal sequences

- $n = |Q|$ is number of states
  $m = |I| * n$ is number of transitions

Mealy machines

- Deterministic
- Input-enabled
- Outputs on transitions
  - (Motivated by testing)



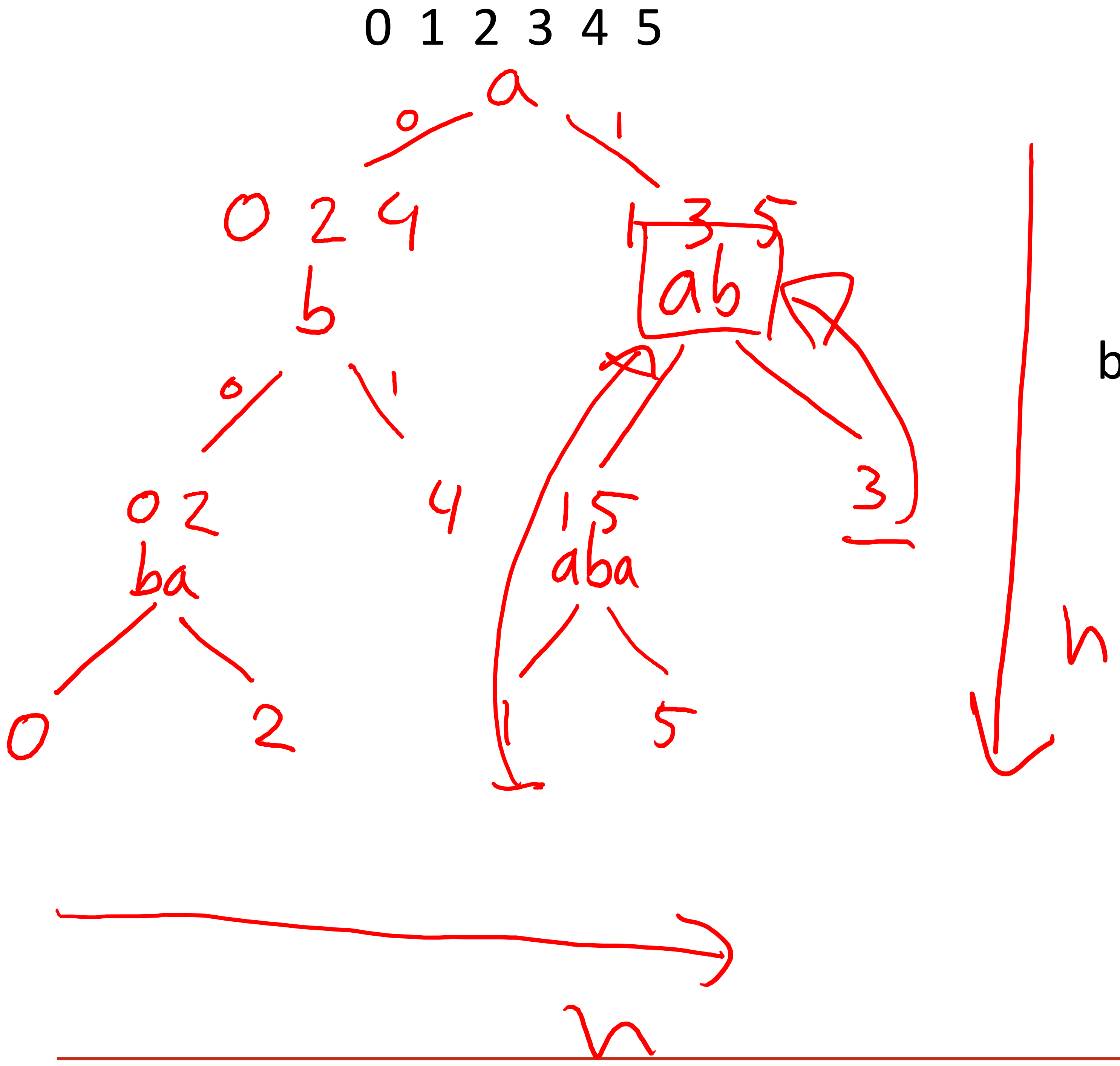
- States s, t are *equivalent* if $[\![s]\!](w) = [\![t]\!](w)$ for all w, where $[\![s]\!]: I^* \to O^*$
- We are interested in *inequivalence!*

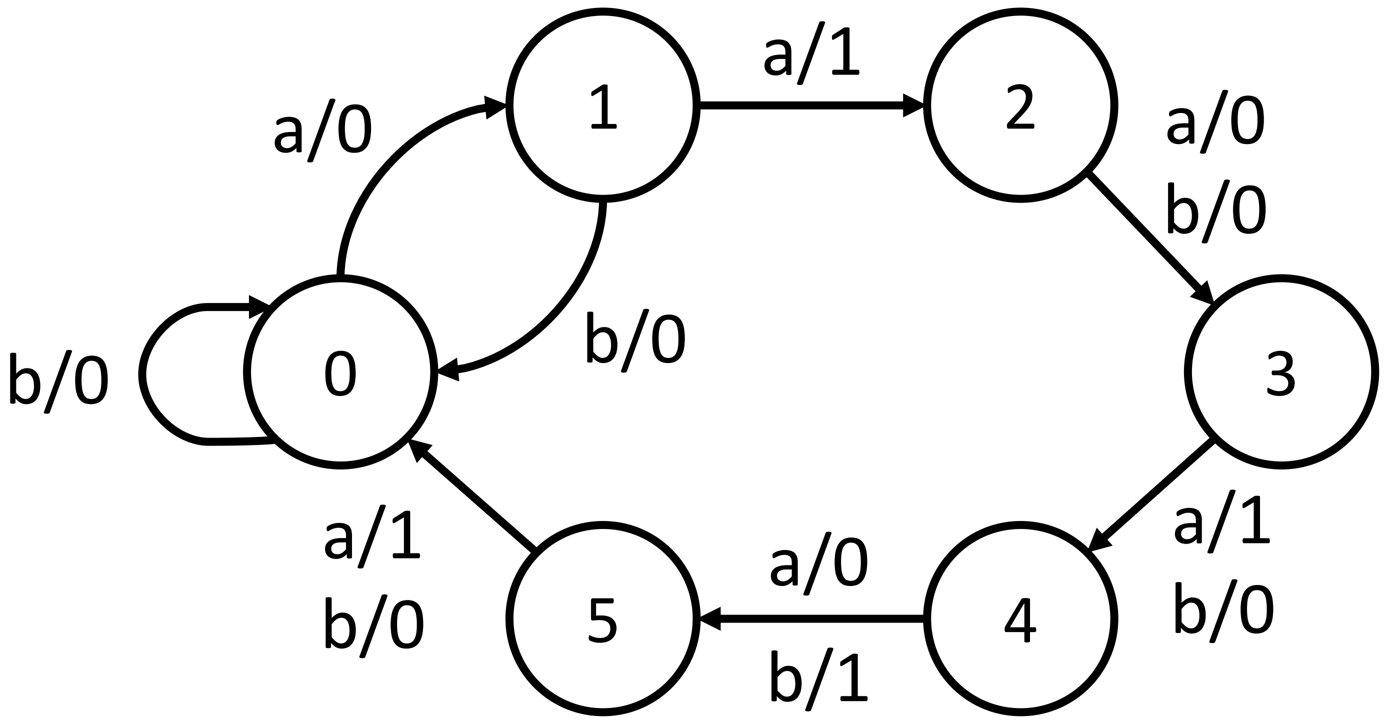
Basic Partition Refinement

Roughly:

- Start with trivial partition
- Split classes if
  1. states have different output, or
  2. states transition to different classes.

Partition Refinement example





Worst-case $O(n^2 * |I|)$

Order of splitting not specified

Key improvements

- Hopcroft's algorithm
  - Keep a queue of splitters
  - Skip the largest set in the splitter
  - (Note that we allow more than two outputs.)

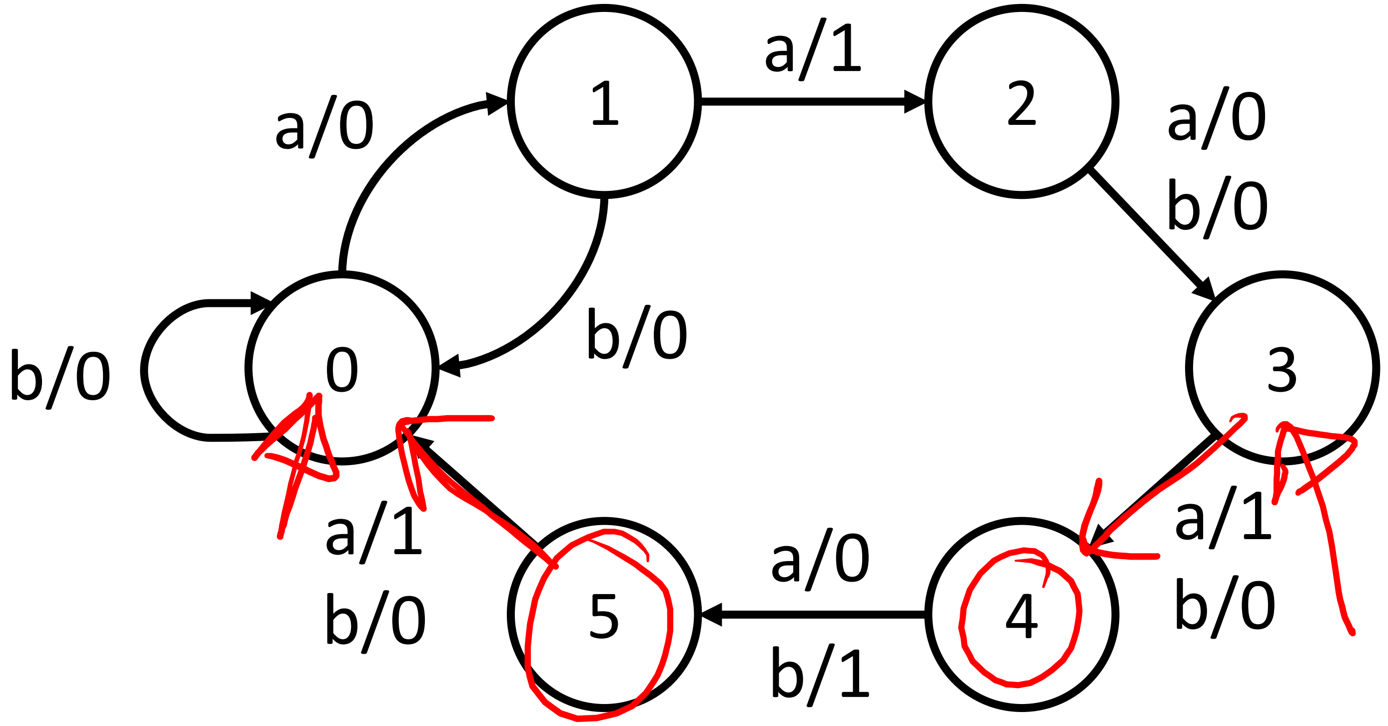
- Minimality:
  - Queue in order of size

- Witnesses:
  - As linked list, copying suffix is too expensive

Hopcroft example

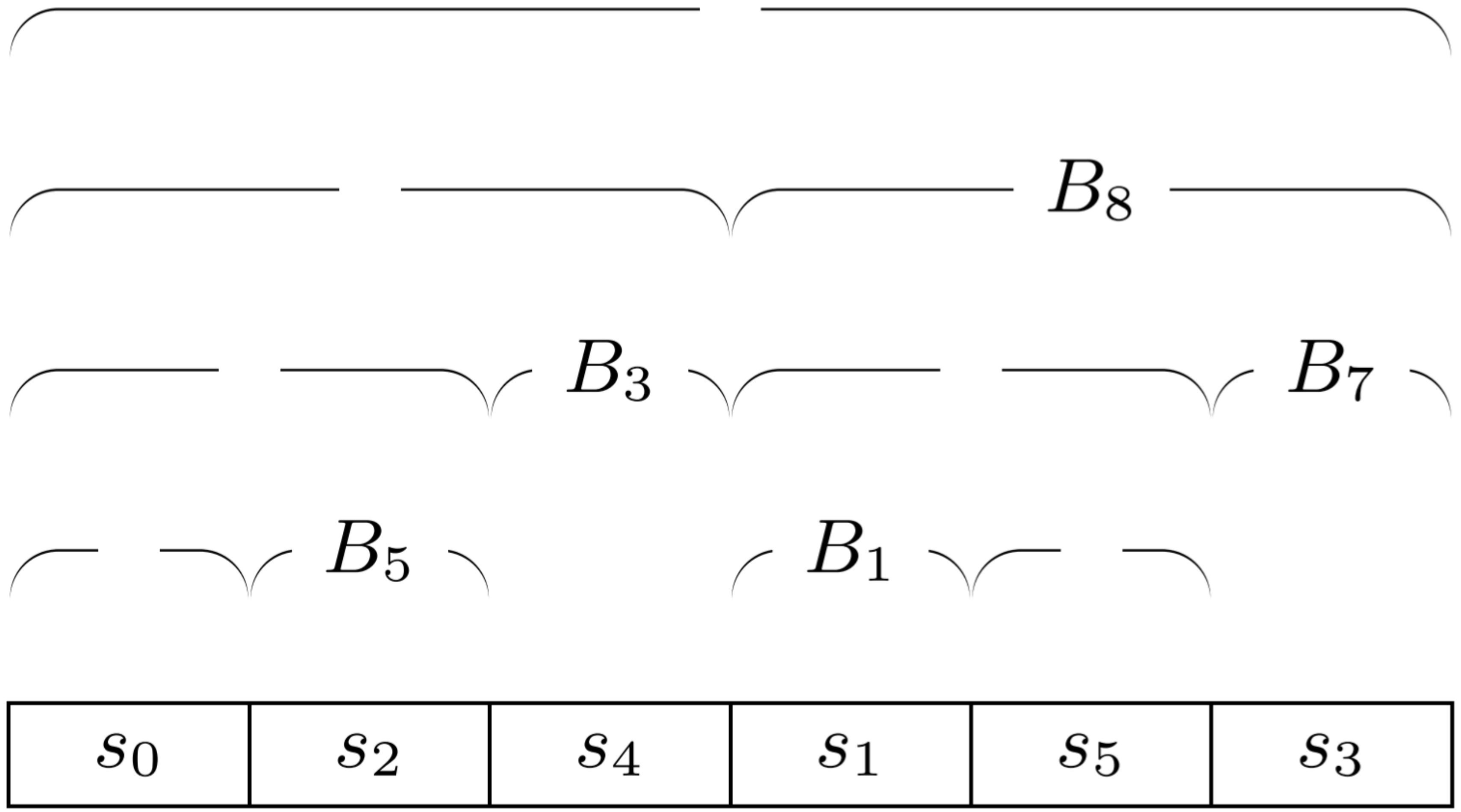
Queue

0 1 2 3 4 5



Why O(n log n)?

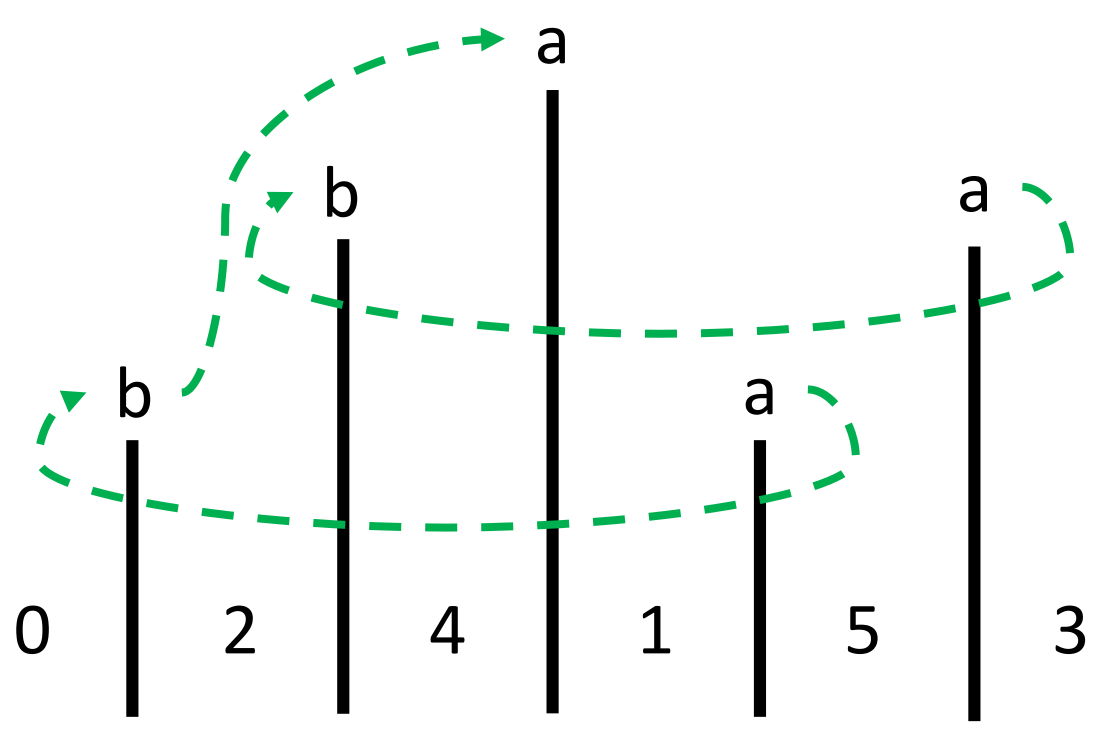
- Let B = nodes which are not the largest
- Every state is in at most $\log n$ elements of B
- => Every state is `touched` at $\log n * |I|$ times
- => Gives $O(m \log n)$ bound

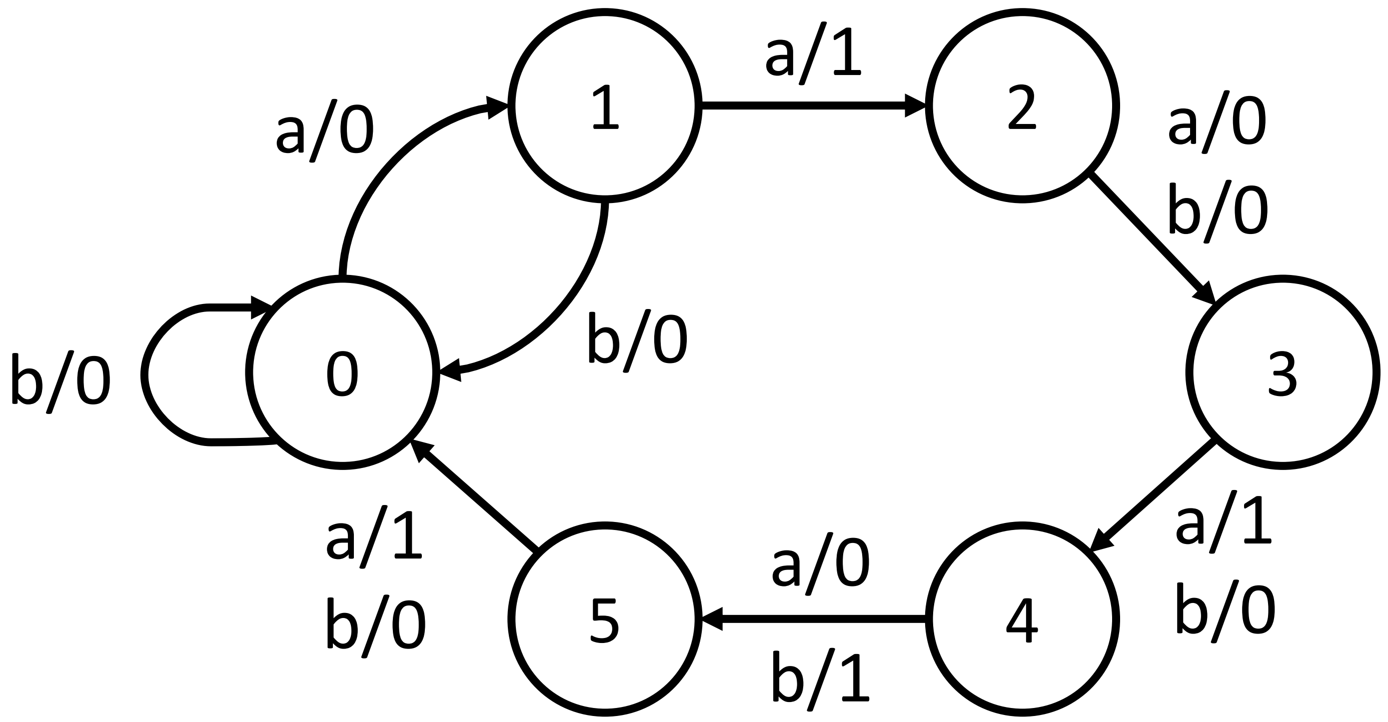


Bookkeeping

- Pre-processing of δ^{-1} in $O(m)$
- Sorting in linear time (a la Dutch flag problem)
- Counters to determine largest child
  and to check whether a node is split
- Separating sequences stored as linked lists

End result





- Small data structure containing minimal separating sequences for all pairs
- Space $O(n)$
- Query time $O(n)$

Application: black-box conformance testing

- Problem:
- Given a specification M and a **black-box** system X (both Mealy machines),
  Can we decide $X \approx M$ by performing an experiment?
- If X is too big, this is impossible, so we ask for an experiment deciding:
  If $|X| \leq |M|$, then $X \approx M$?
- Chow and Vasilevskii (independently) in ~1970 gave a experiment of polynomial size!
- **W-method**: Test suite = $P \cdot W$ ← char. set

↑ State+Transition cover

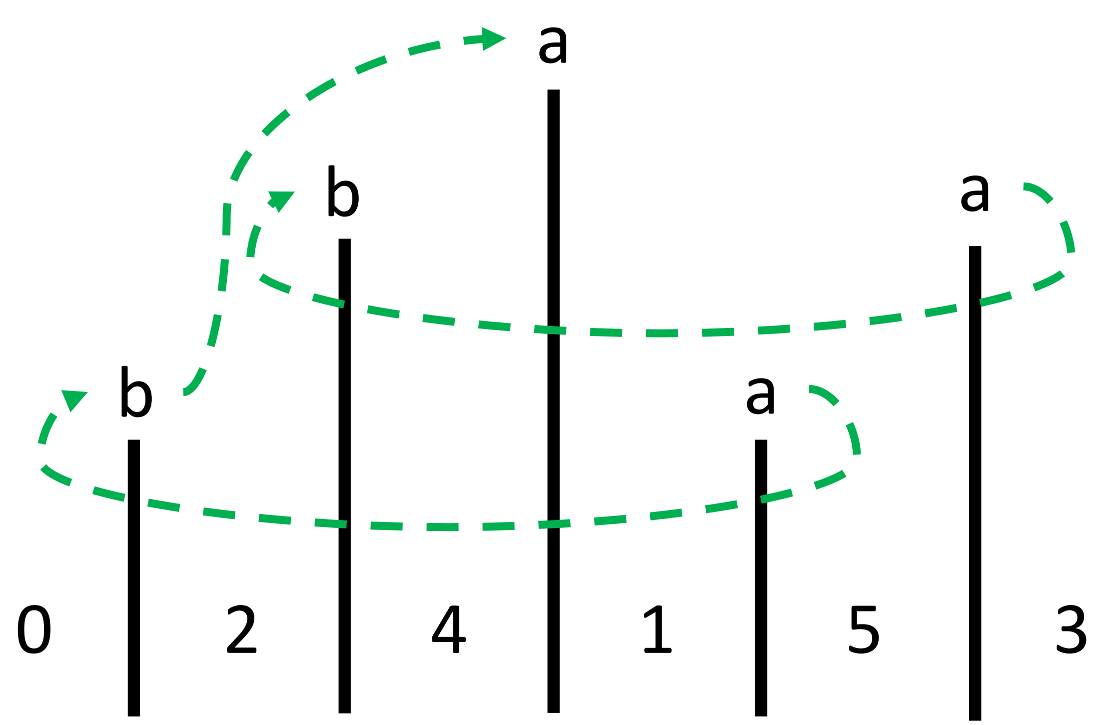
Test suites

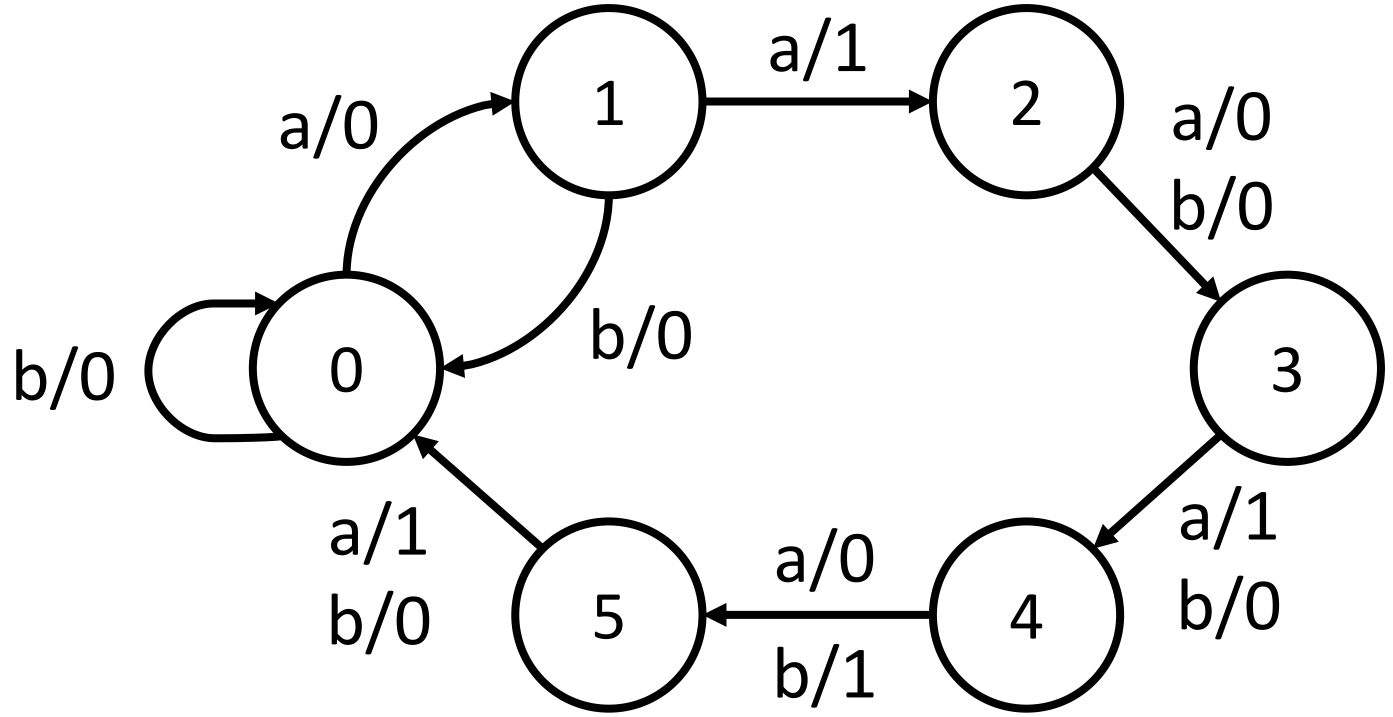
- **W-method**
  Characterisation set = set containing a separating sequence for each pair.
  Constructible in $O(m \log n)$.

- **Wp-method**
  Local state identifier for s = set containing a separating sequence for each other t.
  Constructible in $O(m \log n + n^2)$.

- **HSI-method**
  Same as Wp-method, but requires state identifiers to be *harmonised*. Our construction guarantees this.
  Now $O(m \log n + n^2)$, previous $O(mn^3)$.

- Typically we remove common prefixes: $O(|W|)$ or $O(n^2)$.





Minimal Separating Sequences for All Pairs of States in $O(m \log n)$

Rick Smetsers **Joshua Moerman** David N. Jansen

Thanks for your attention!