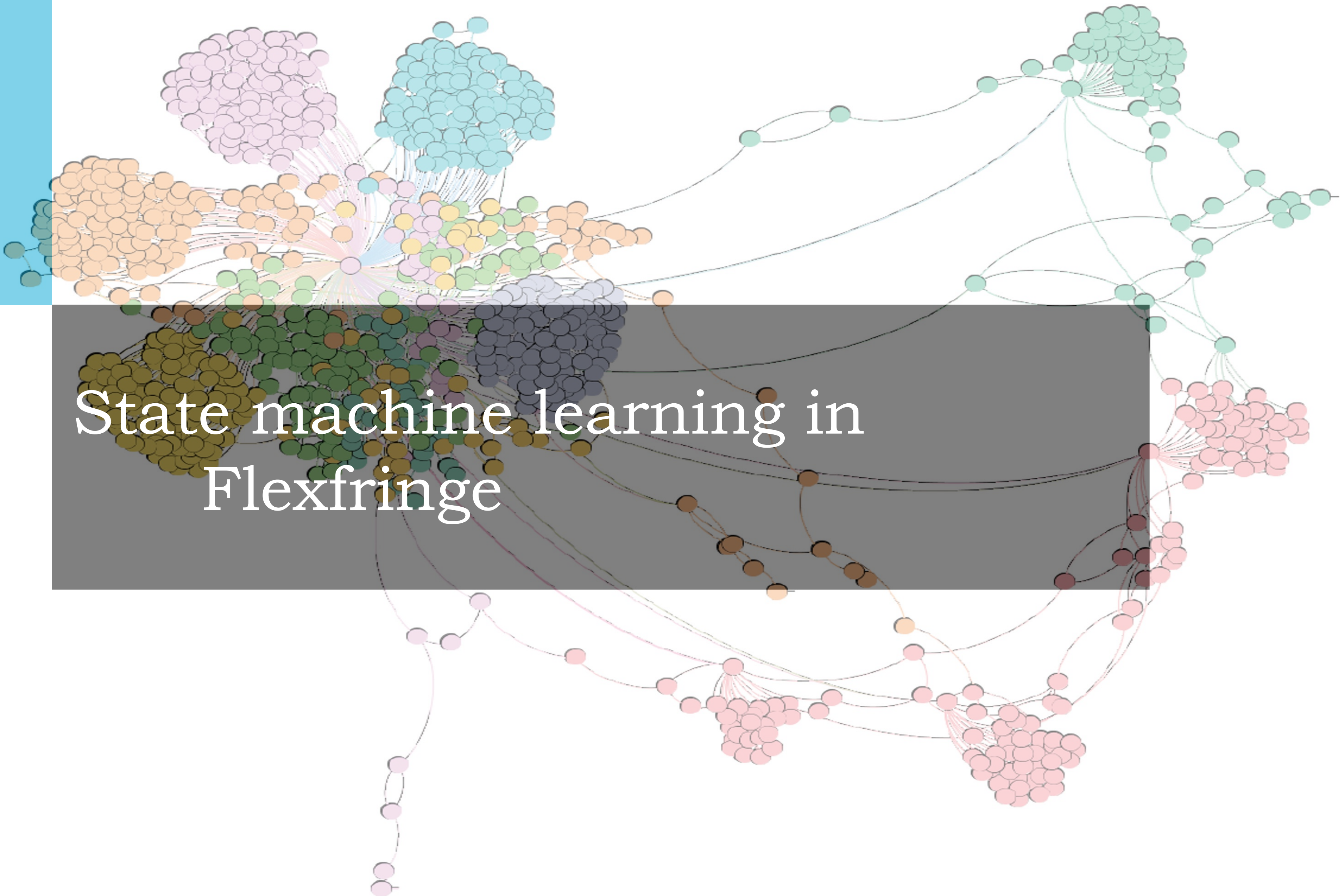
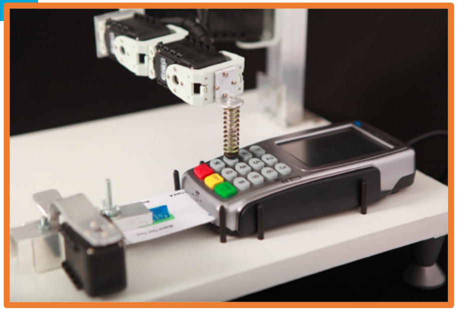
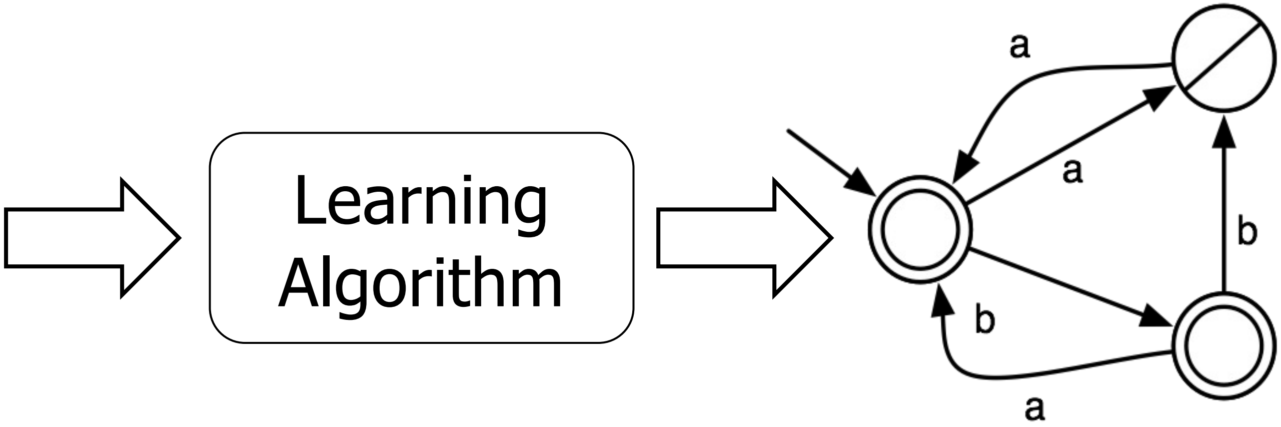


State machine learning in Flexfringe

Cyber Analytics Lab
TU Delft
Sicco Verwer - s.e.verwer@tudelft.nl

Passive Model Learning

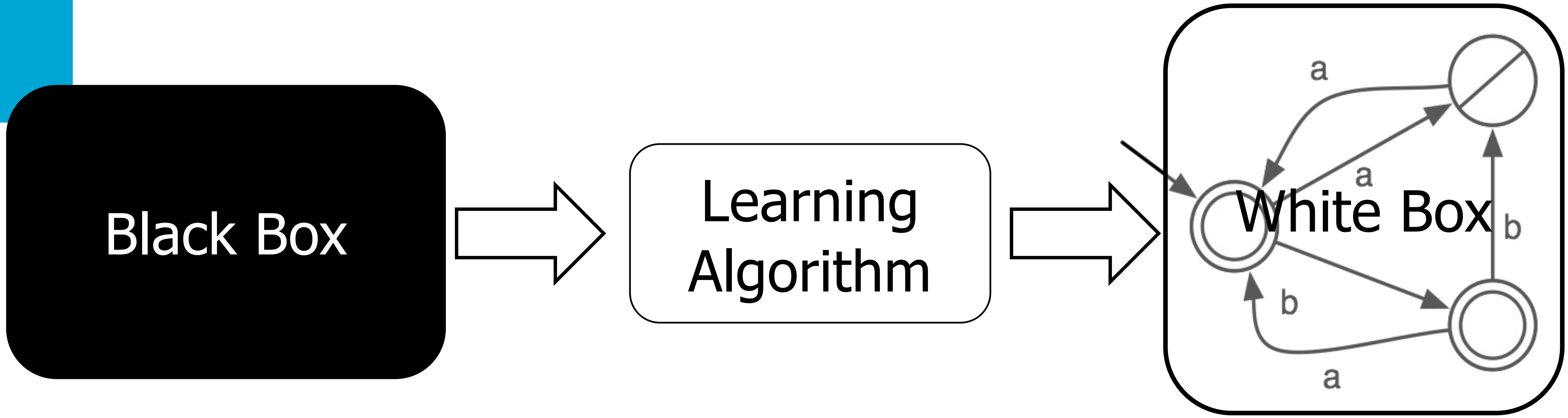




Learning Algorithm

- Software leaves traces (logs)
- A state machine is a logical model describing these traces
  - Classification – is a new trace generated by the same software?
  - Prediction – what trace is most likely to occur next?
  - Analysis – is the software deadlock-free, secure, malicious?

Passive Model Learning



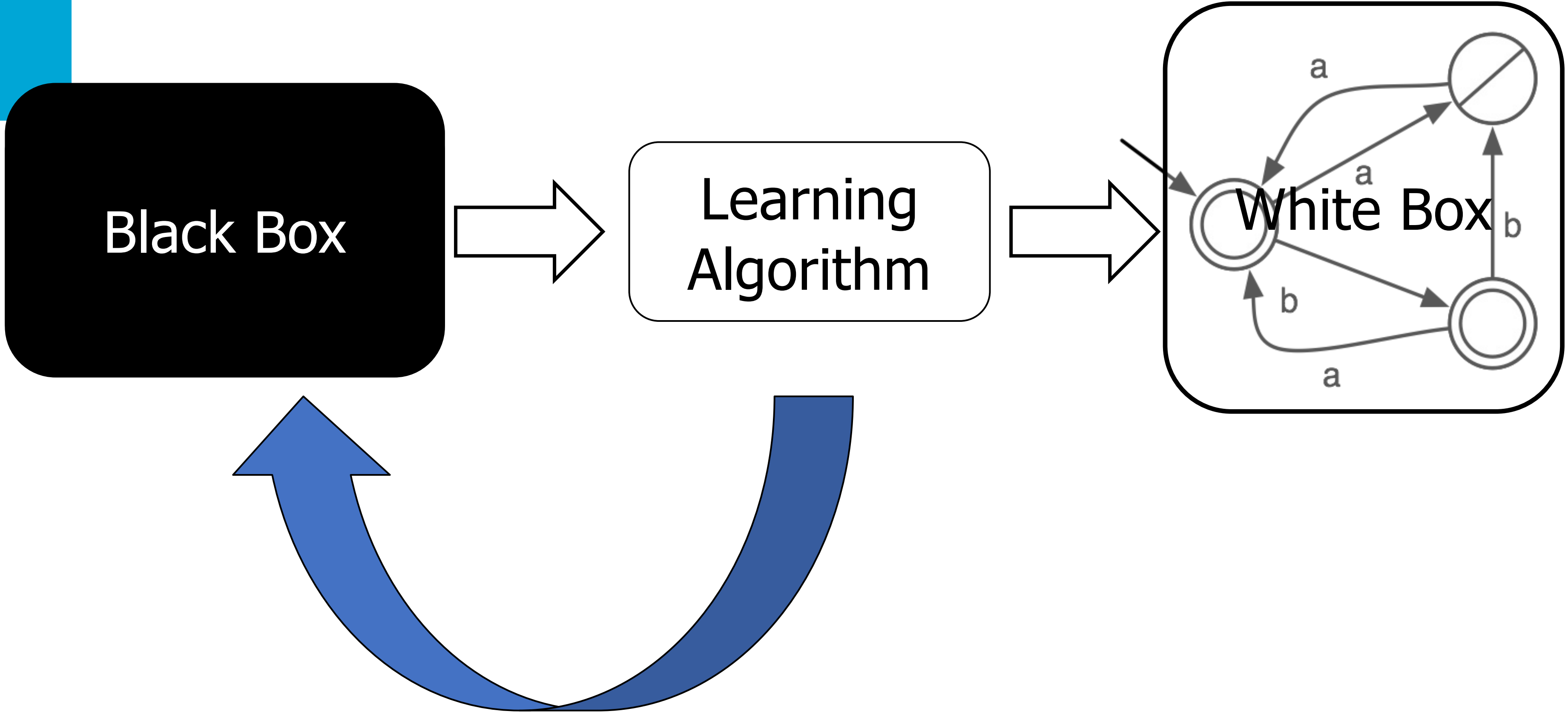
Black Box

Learning Algorithm

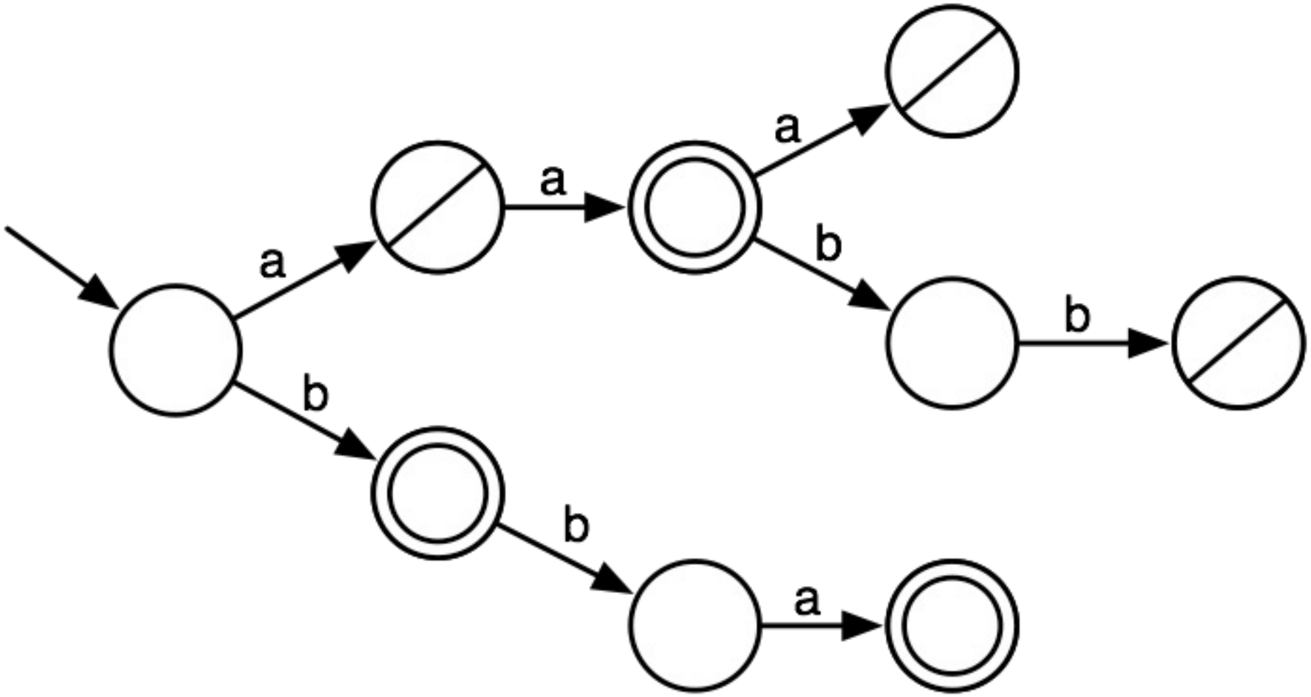
White Box

- Software leaves traces (logs)
- A state machine is a logical model describing these traces
  - Classification – is a new trace generated by the same software?
  - Prediction – what trace is most likely to occur next?
  - Analysis – is the software deadlock-free, secure, malicious?

Active Model Learning

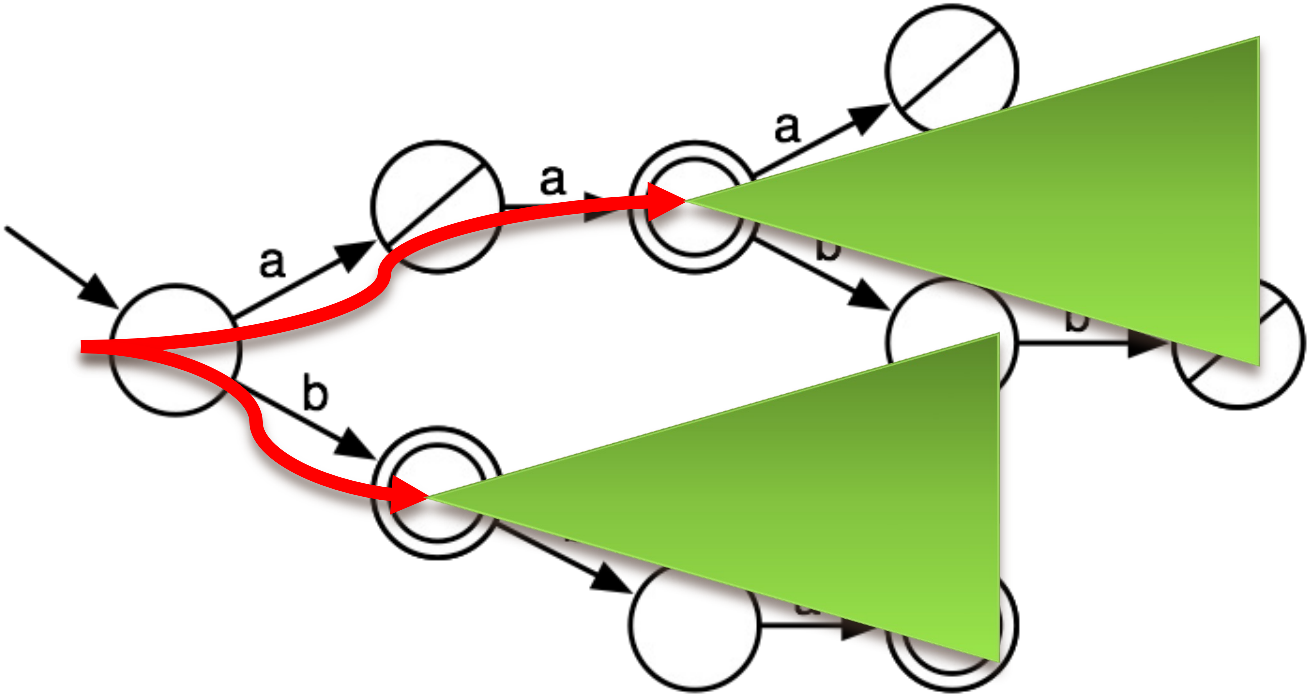


Learning DFAs



positive data: aa, b, bba; negative data: a, aaa, aabb
represented as a prefix tree

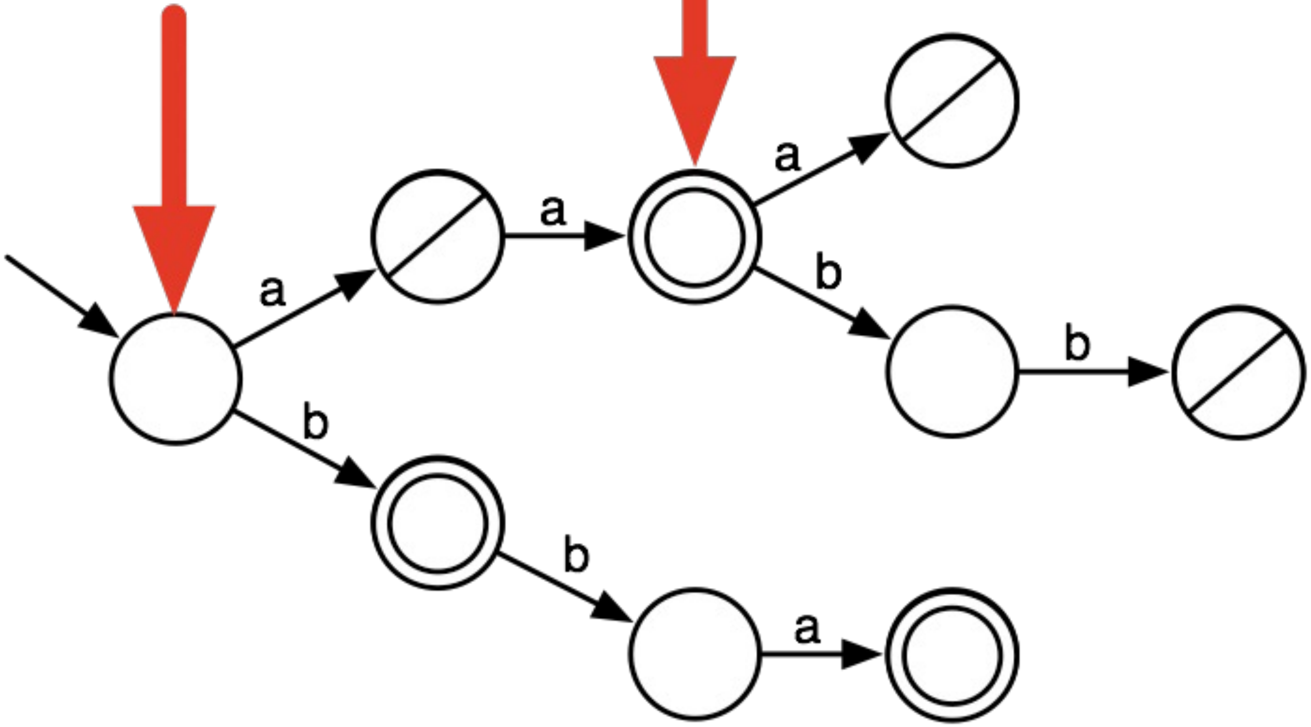
Learning DFAs



Now we test for **Myhill-Nerode or Markov**:

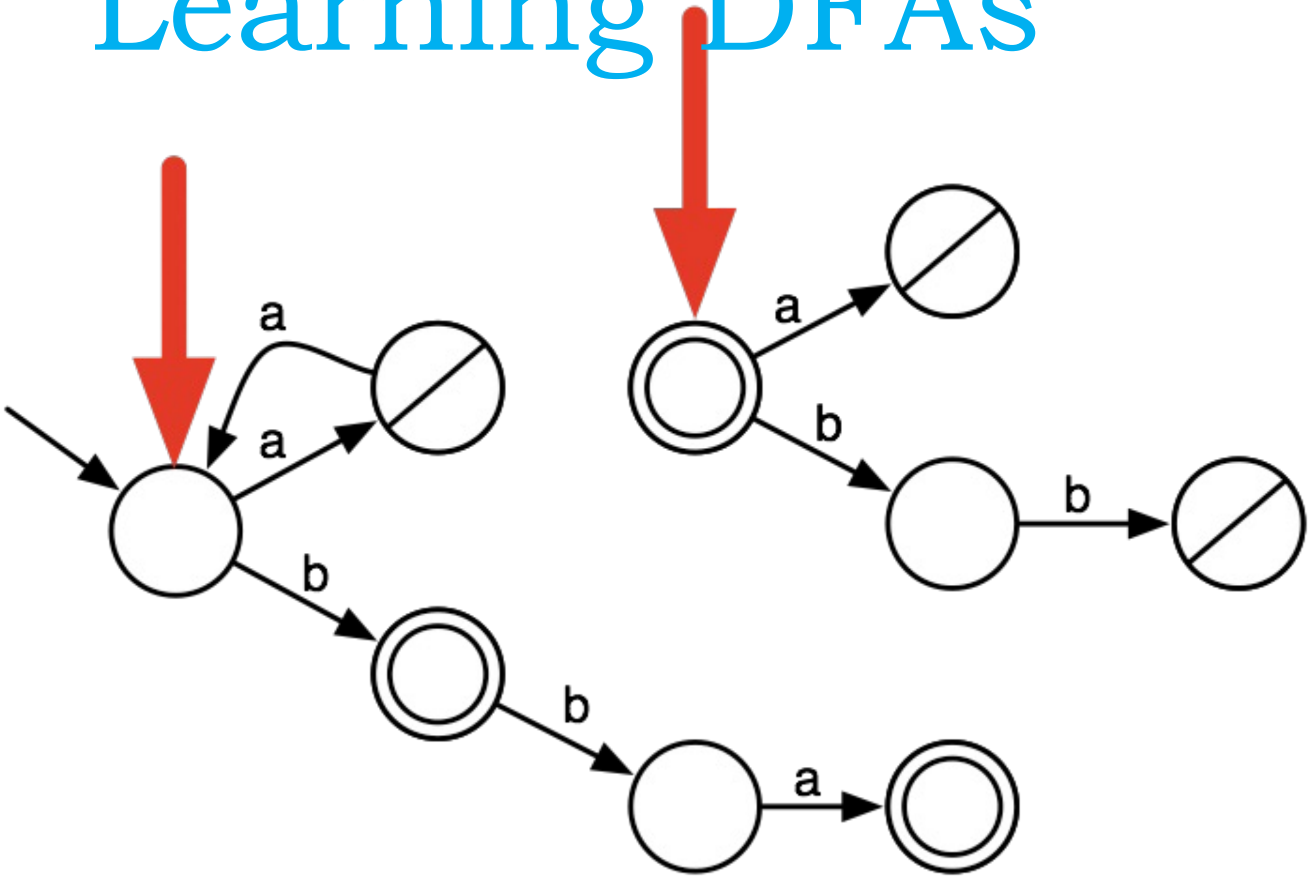
Two states q and q’ are equivalent iff *their future is independent from their past*

Learning DFAs



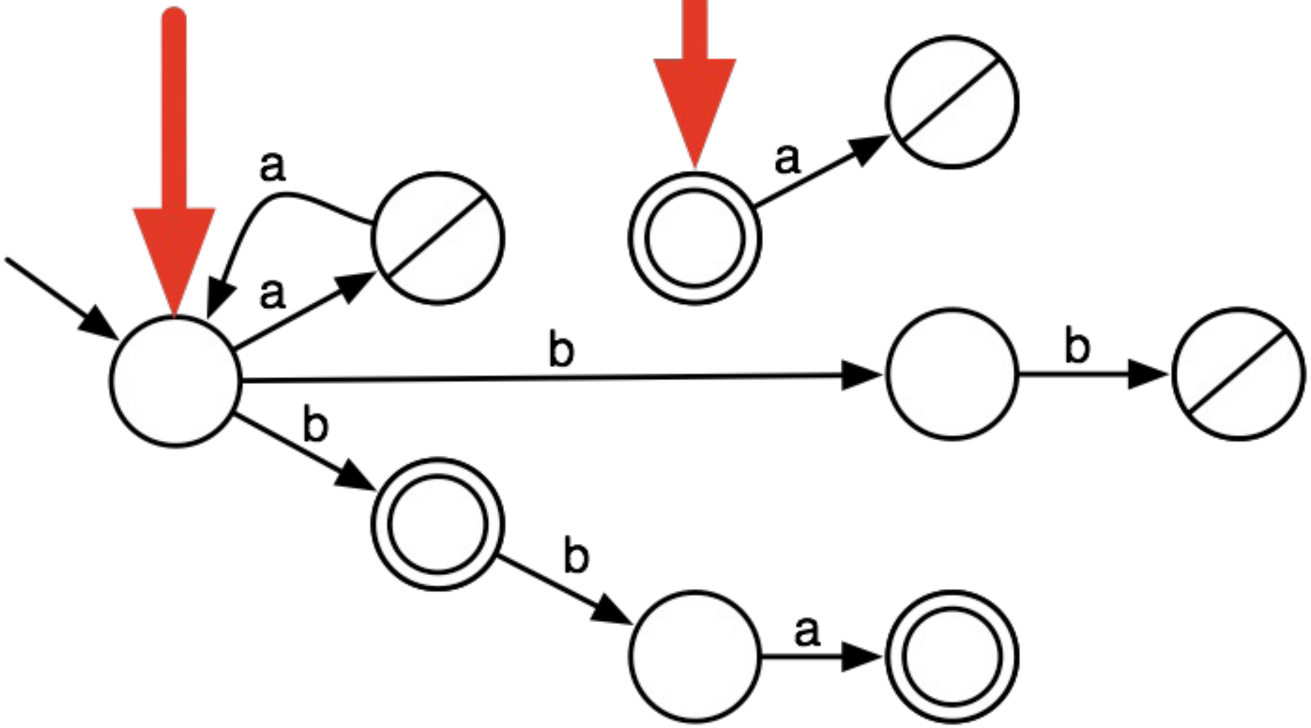
State merging:
select two nodes

Learning DFAs



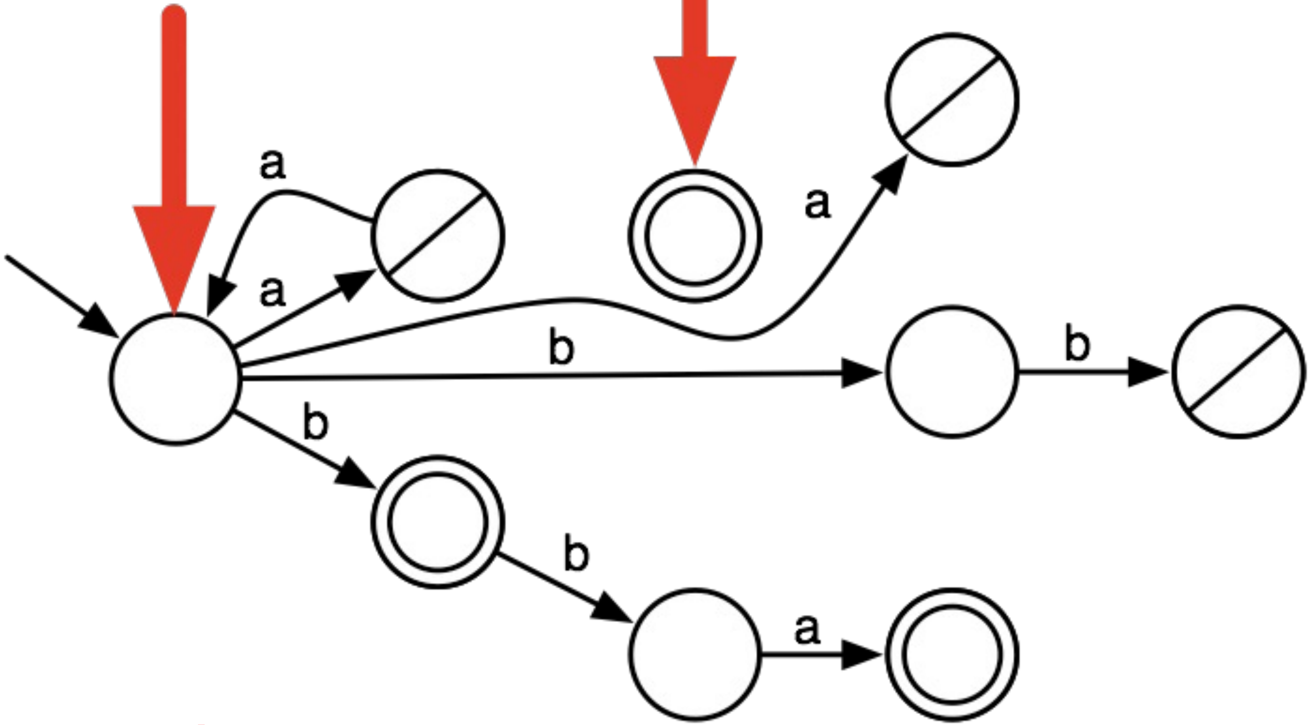
State merging:
move input transitions from one state to the other

Learning DFAs



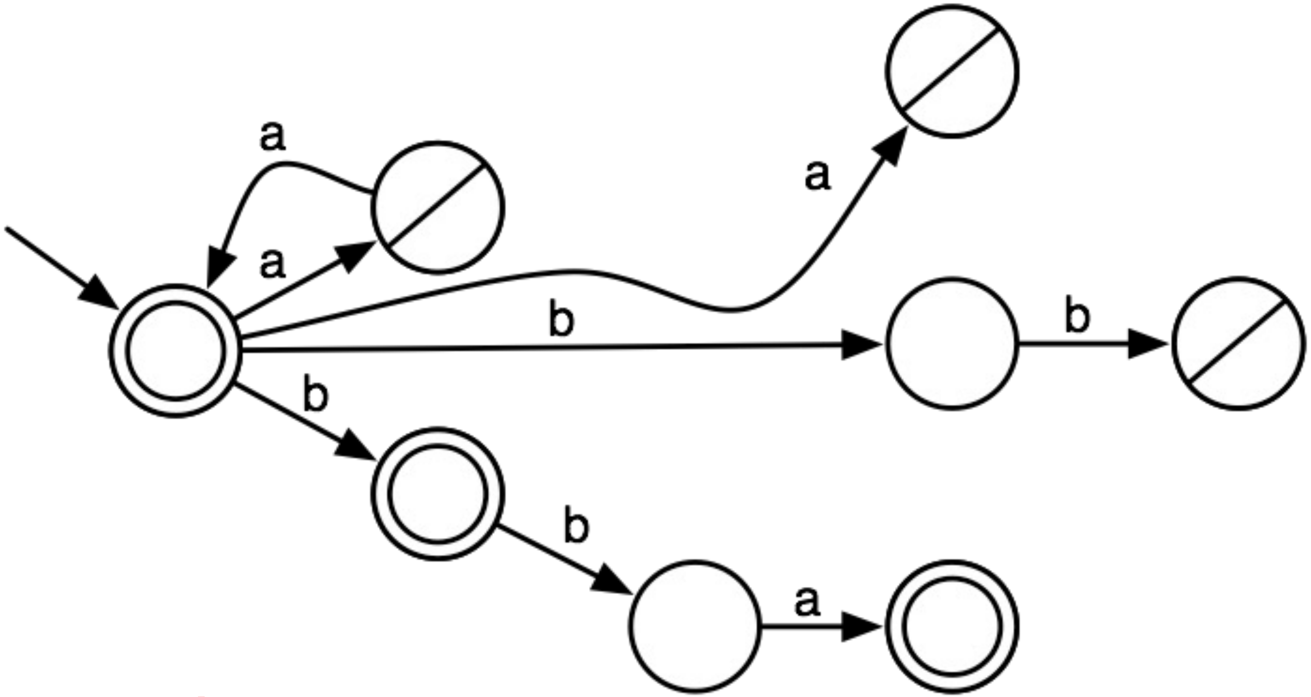
State merging:
move input transitions from one state to the other

Learning DFAs



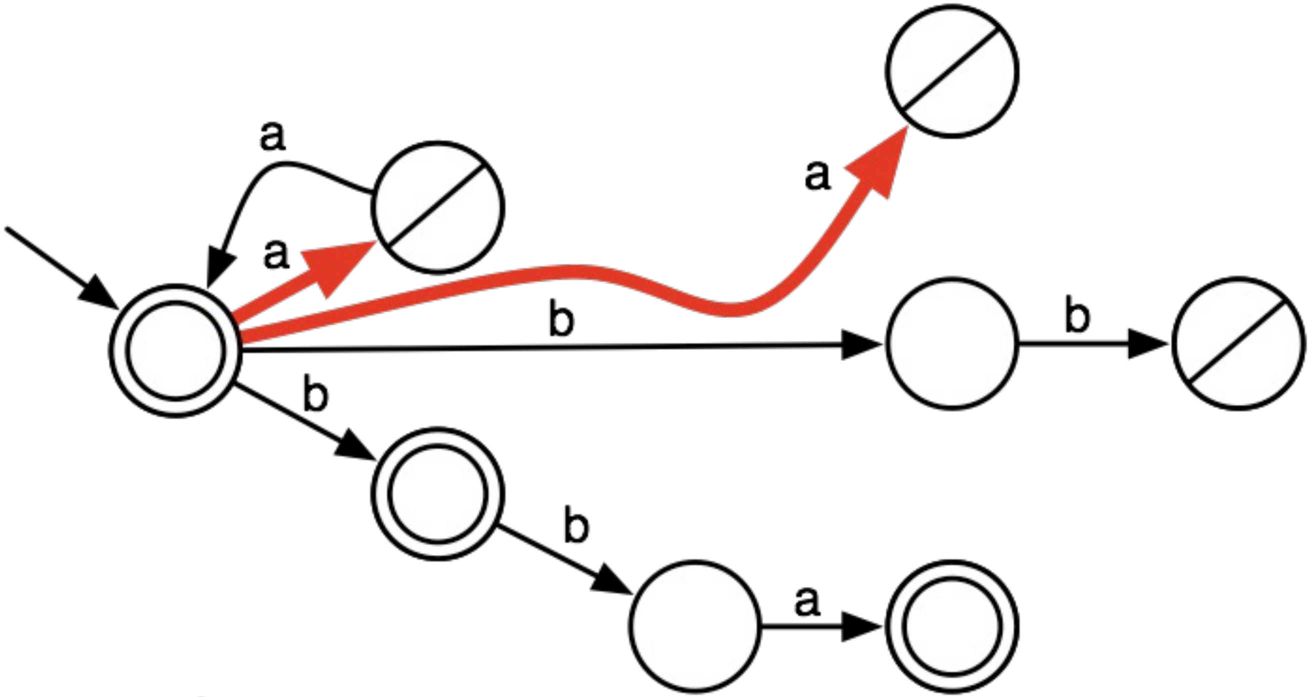
State merging:
move output transitions

Learning DFAs



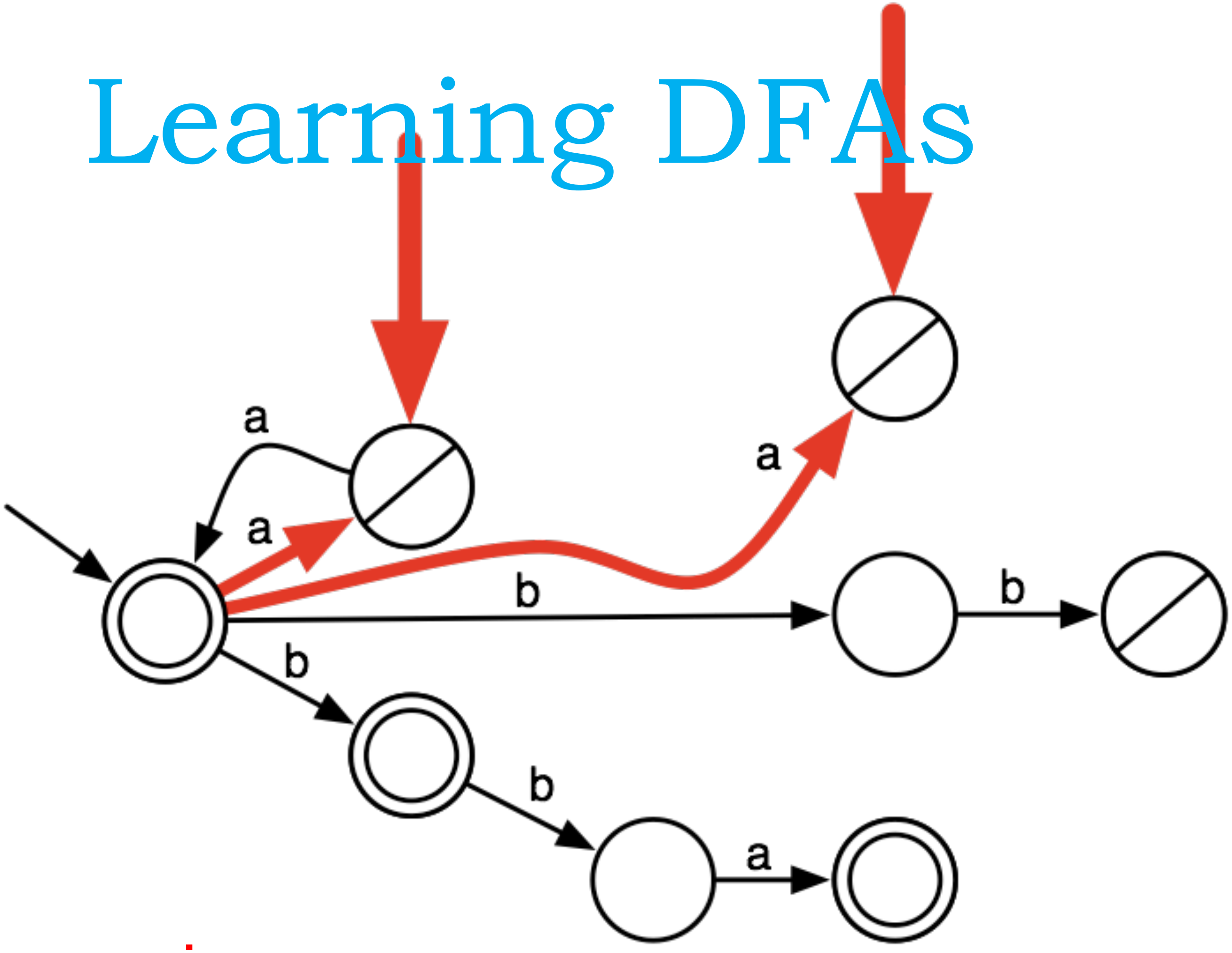
State merging:
delete the obsolete state, maintain pos/neg

Learning DFAs



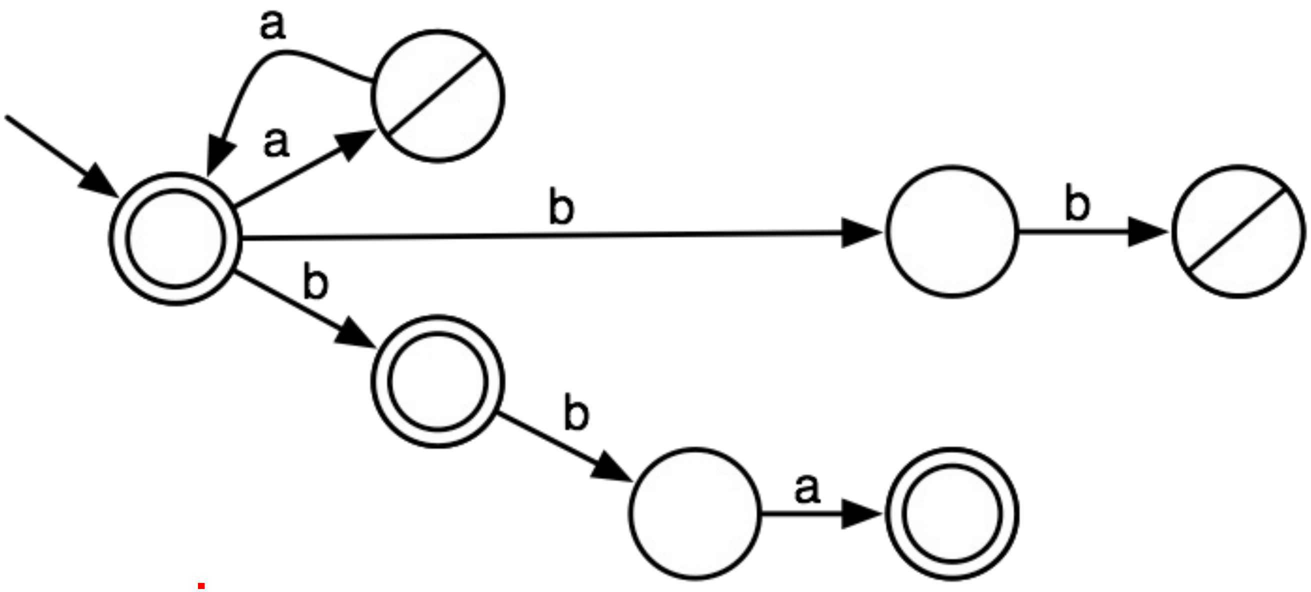
State merging:
merge targets of non-deterministic transitions

Learning DFAs



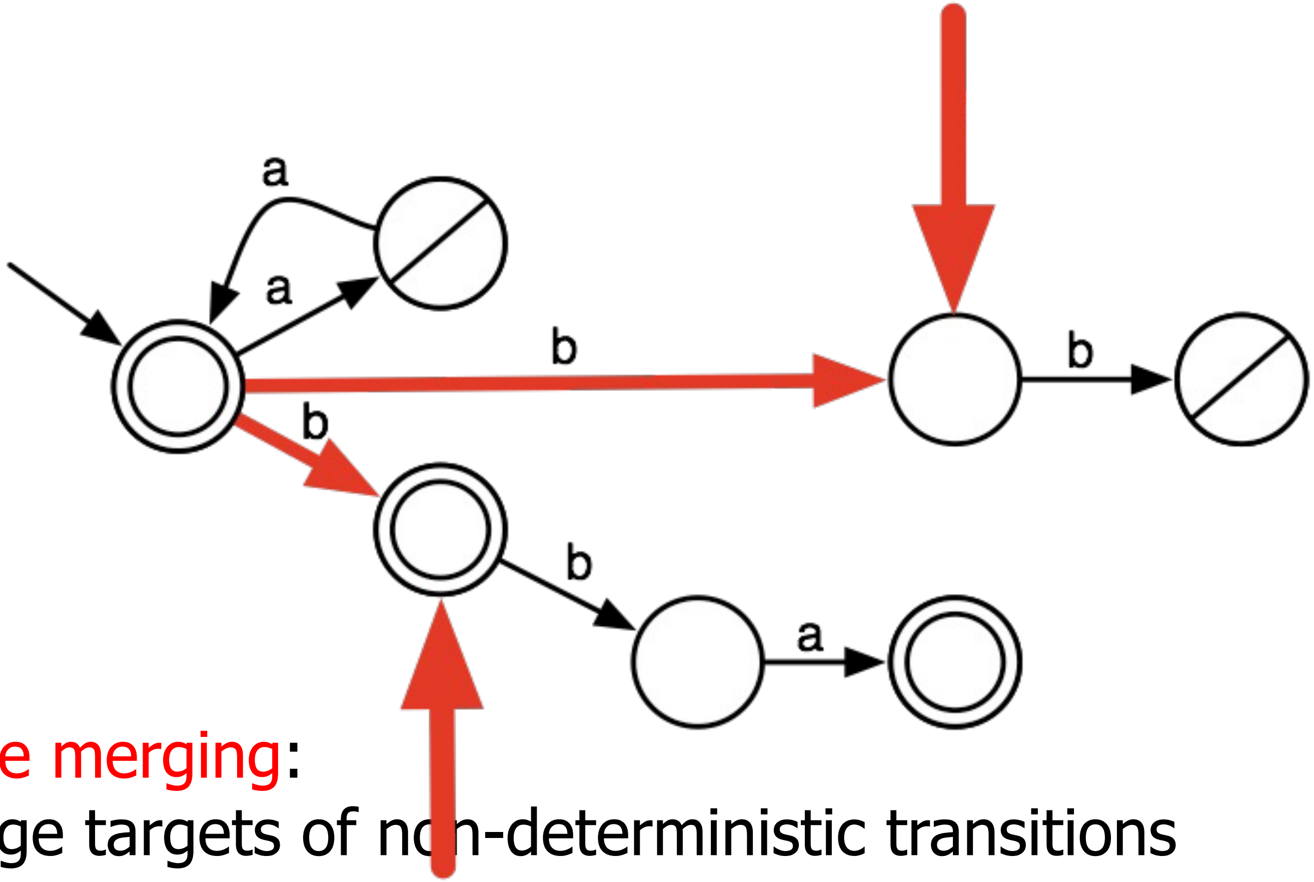
State merging:
merge targets of non-deterministic transitions

Learning DFAs



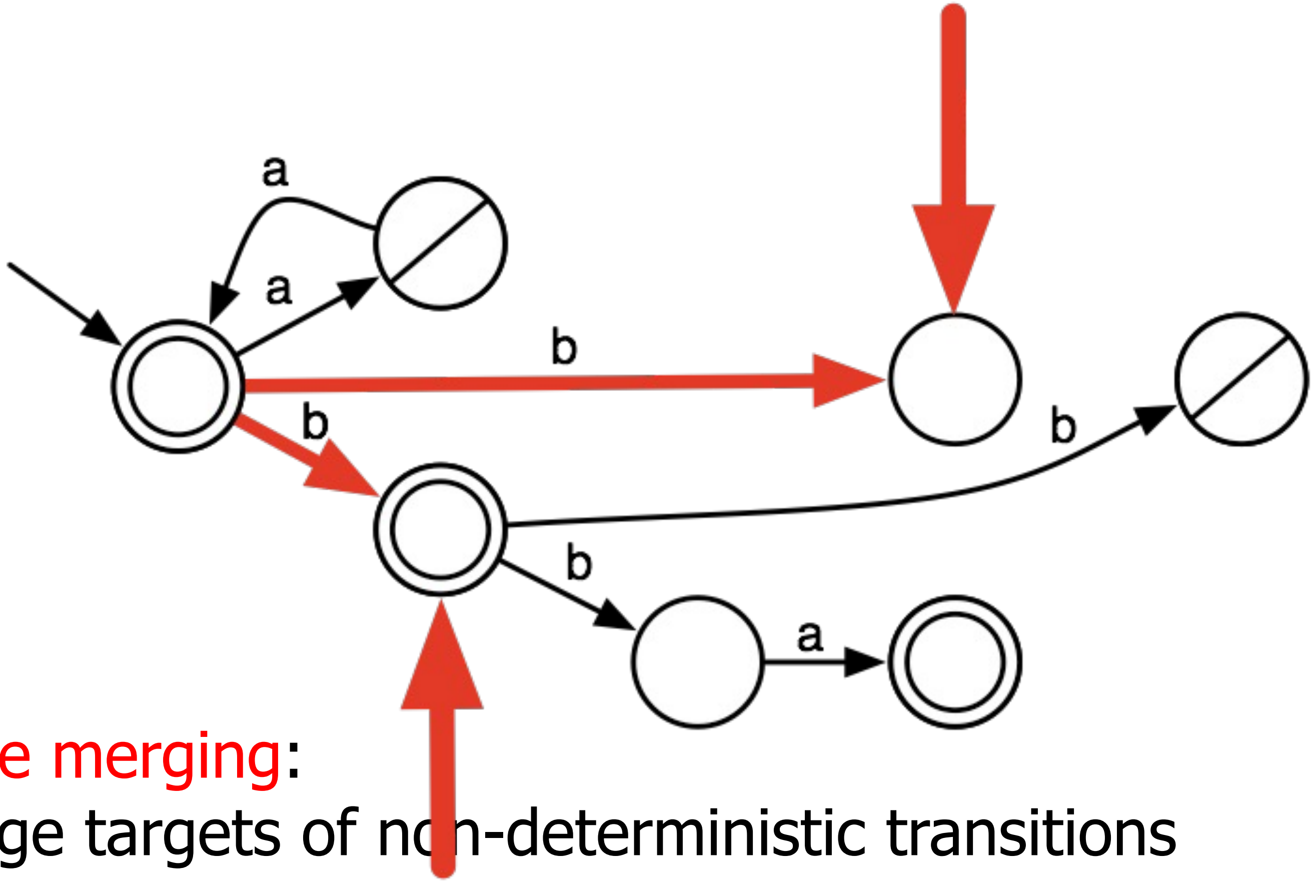
State merging:
merge targets of non-deterministic transitions

Learning DFAs



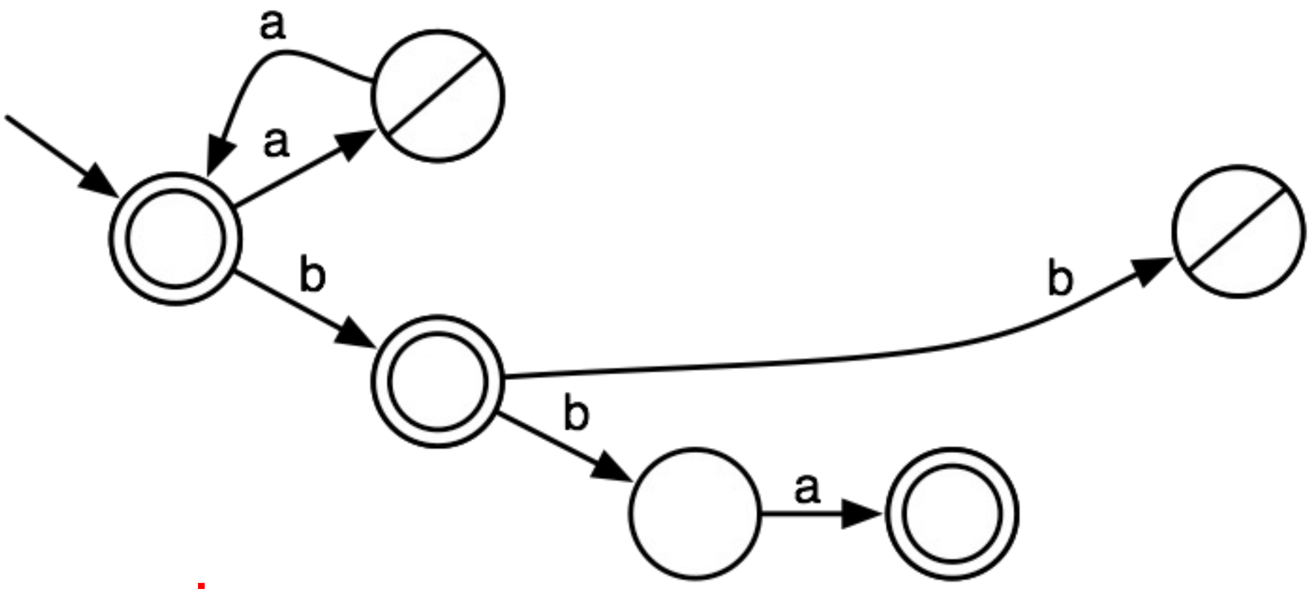
State merging:
merge targets of non-deterministic transitions

Learning DFAs



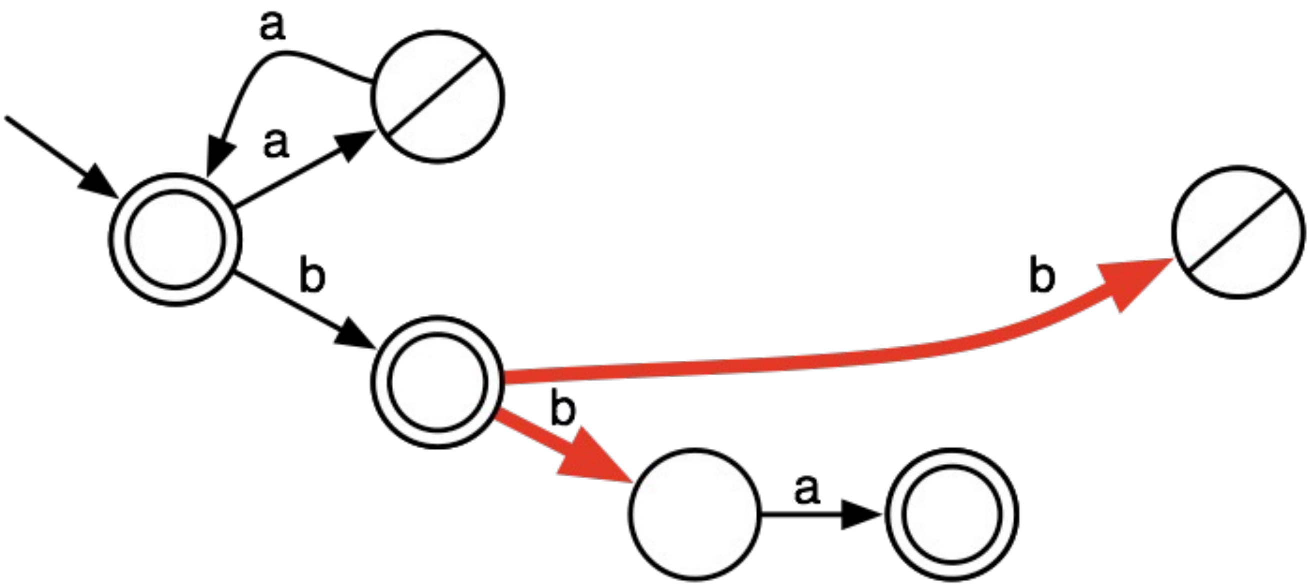
State merging:
merge targets of non-deterministic transitions

Learning DFAs



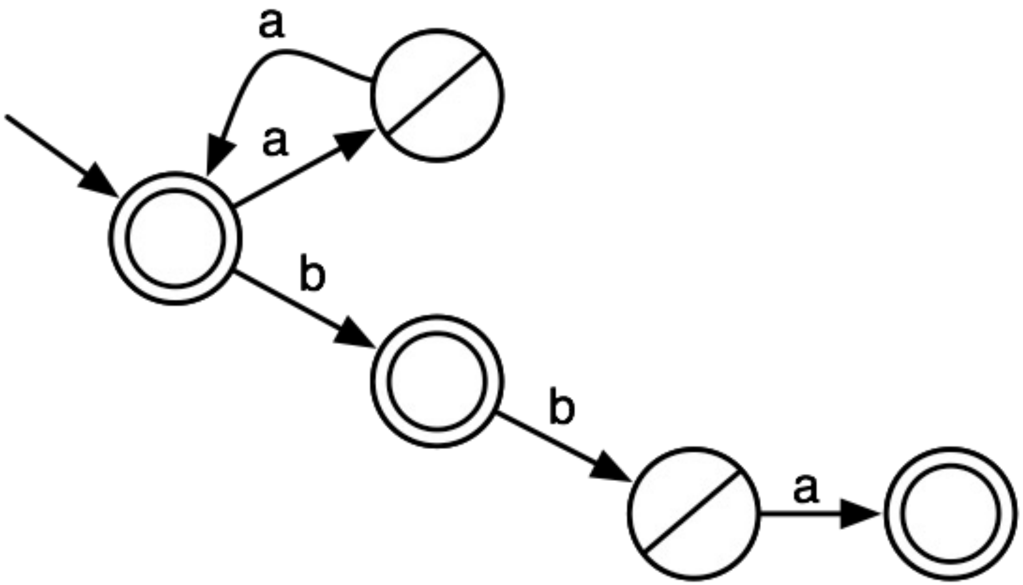
State merging:
merge targets of non-deterministic transitions

Learning DFAs



Select two new nodes to merge and iterate

Learning DFAs

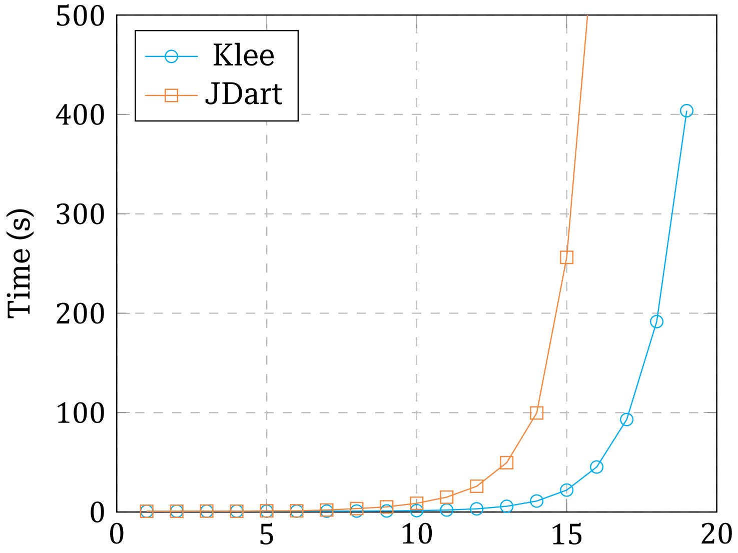


Select two new nodes to merge and iterate

Application: fuzzing loops

```
int main(int argc, char** args) {
  assert(argc == 3);
  int limit = atoi(args[1]); // Target to reach
  int i = 0; // Internal state
  int j = 0; // Loop variable
  char symbol; // Character in input
  char* trace = args[2]; // Input: array of symbols

  while ((symbol = trace[j++]) != 0) { // Get next character in input
    if (symbol == 'i') {
      i += 1;
    } else if (symbol == 'p') {
      if (i >= limit){
        assert(0); // Crash
      }
    } else {
      return 0;
    }
  }
  return 0;
}
```

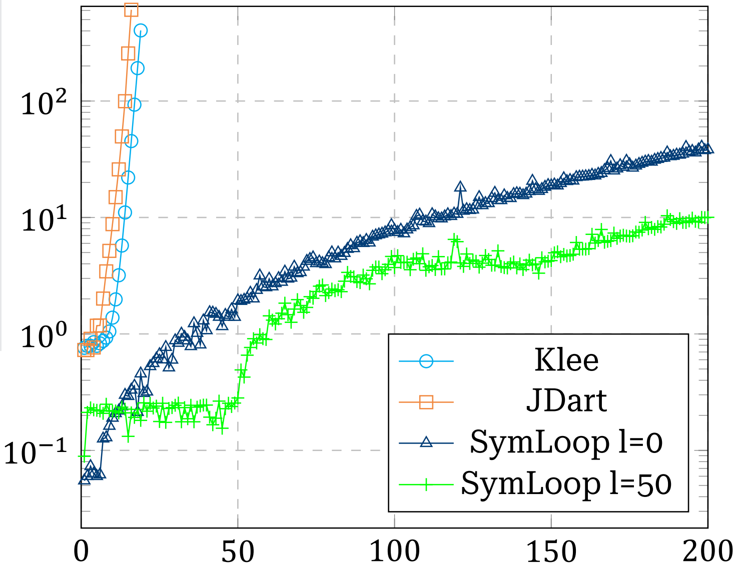


with Bram Verboom and Simon Dieck

Application: fuzzing loops

```
int main(int argc, char** args) {
  assert(argc == 3);
  int limit = atoi(args[1]); // Target to reach
  int i = 0; // Internal state
  int j = 0; // Loop variable
  char symbol; // Character in input
  char* trace = args[2]; // Input: array of symbols

  while ((symbol = trace[j++]) != 0) { // Get next character in input
    if (symbol == ’i’) {
      i += 1;
    } else if (symbol == ’p’) {
      if (i >= limit){
        assert(0); // Crash
      }
    } else {
      return 0;
    }
  }
  return 0;
}
```



with Bram Verboom and Simon Dieck

Learning from intrusion data

From a learned model,

we extract all paths leading

to severe objectives

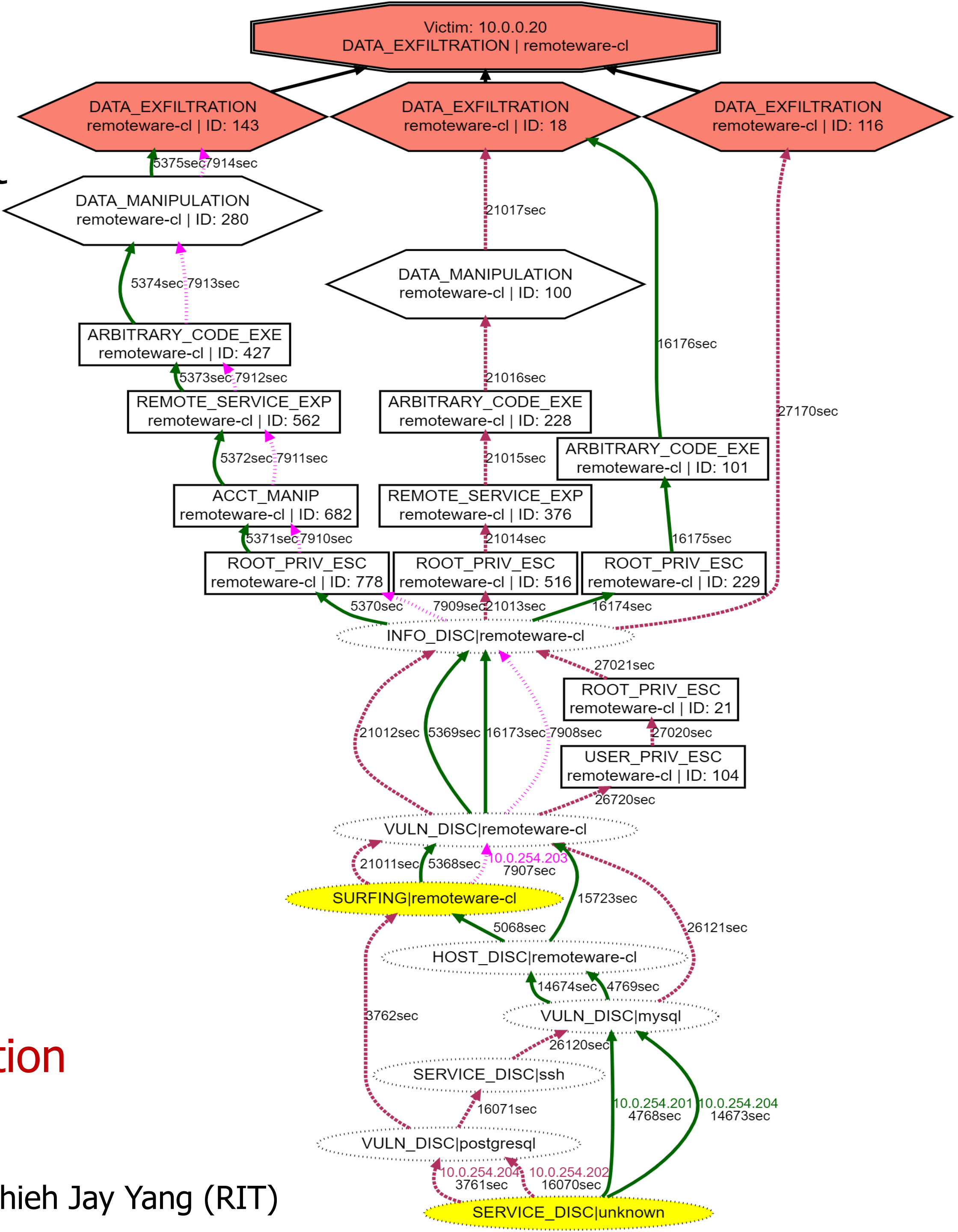
Paths are time-stamped and

colored per attacker

Right:

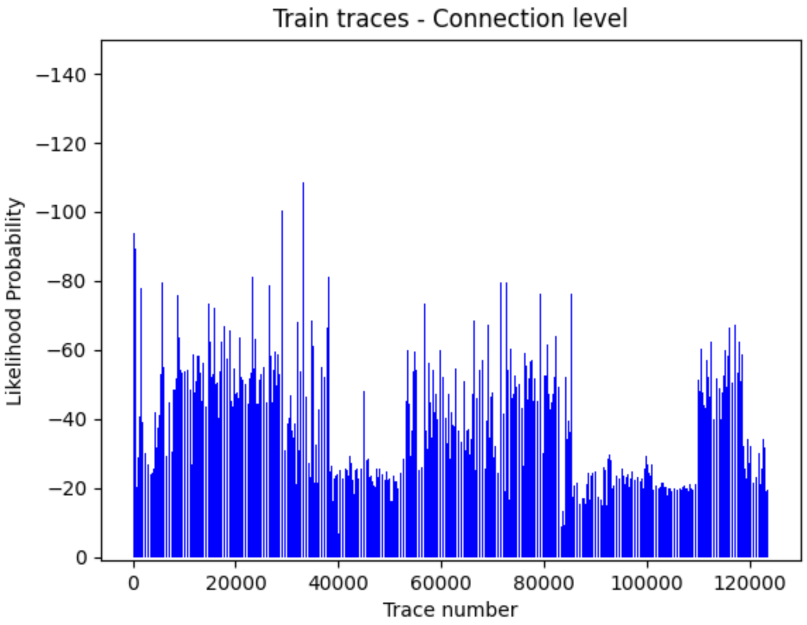
3 teams showing different

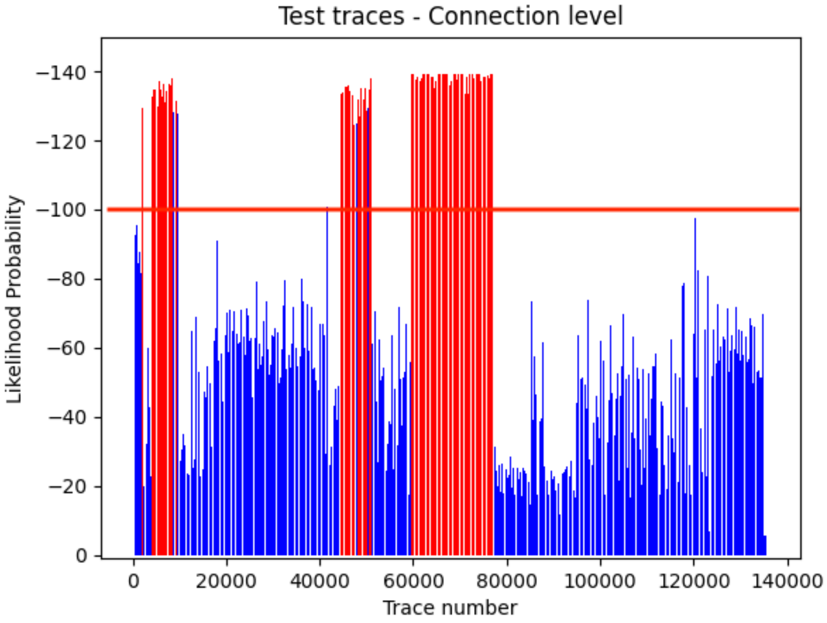
ways to reach data exfiltration



with Azqa Nadeem, Stephen Moskal and Shanchieh Jay Yang (RIT)

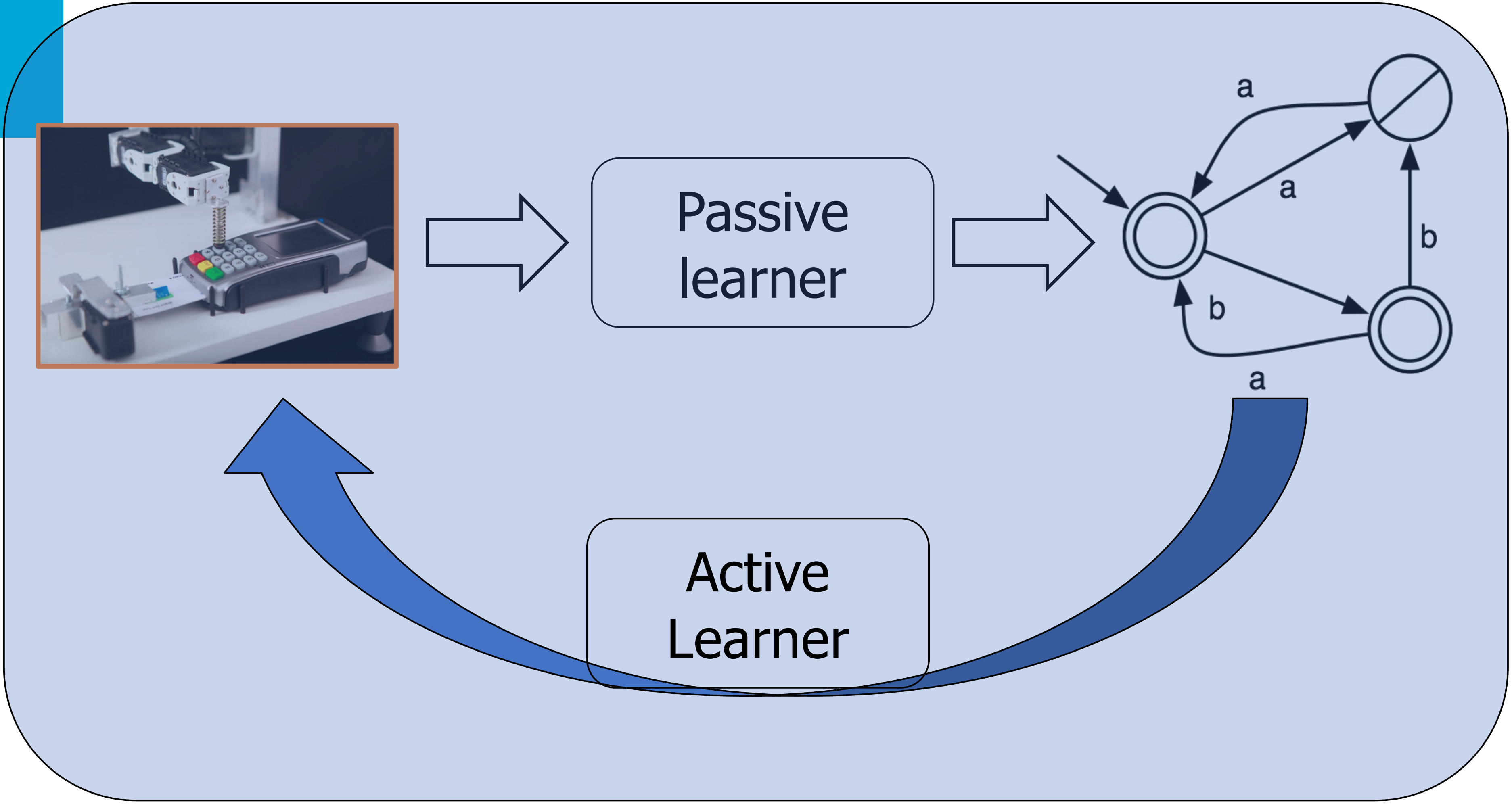
Detection intrusions from NetFlow





With Clinton Cao

Active? Work in progress



Passive learner

Active Learner