The problem of converting a deterministic finite automaton(DFA) to a minimal unambiguous finite automata(UFA) is NP-COMPLETE

Kecheng Yang

Quick Outline

- Presentation of the problem and Intro
- DFA→UFA is in NP
- DFA→UFA is NP-complete

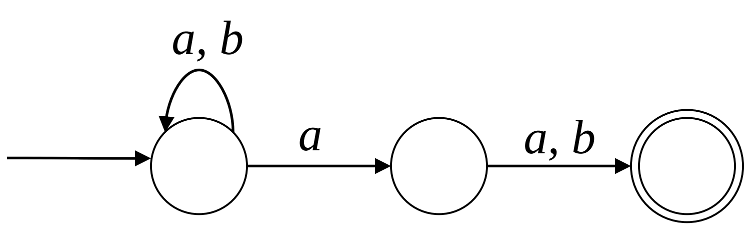
Problem&

DFA → UFA

Given a DFA, find a *minimum* equivalent UFA

The UFA minimization is NP-complete

A Unambiguous FA is an NFA in which every accepted string has a unique accepting computation[Mandel&Simon]



3-state UFA for $\Sigma^* a \Sigma$ with $\Sigma = \{a, b\}$

DFA→UFA is in NP

- 1. There is a deterministic polynomial-time algorithm for deciding whether the two given UFA, M_1 and M_2 are equivalent.
  - Given M_1 and M_2, determine whether L(M_1)⊆L(M_2)
  - Reduce to give as input two UFA's M_3, M_4, such that L(M_3)⊆L(M_4)

- 2. There is a polynomial-time algorithm that, given an NFA M as input, decides whether M is unambiguous

  - Build an NFA M' that accepts L', where L' is the set of strings that can be derived by at least two different accepting paths. Can be done in polynomial time.
  - M is unambiguous iff L(M') is empty.

DFA → UFA is NP-complete

- The vertex Cover problem is reduced to the normal set basis problem. The Normal set basis problem is reduced to the DFA → UFA problem.
- Let C and B be collections of sets. B is said to be a normal basis of C if for each $c \in$ C there is a pairwise disjoint subcollection of B whose union is exactly c.

  Problem: The normal set basis problem is Np-complete

  Proof [Stockmeyer]:
- Let G=(V,E), k be an instance of the problem, where $V = (v_1, v_2, \cdots v_n)$.
- For each v_1, let $c_i = \{x_i, y_i\}$, i= $1,2, \ldots, n$. $\{v_i, v_j\}$ be in E with i<j, we define

$$c_{i,j}^1 = \{x_i, a_{i,j}, b_{i,j}\},$$
$$c_{i,j}^2 = \{y_j, b_{i,j}, d_{i,j}\},$$
$$c_{i,j}^3 = \{y_i, d_{i,j}, e_{i,j}\},$$
$$c_{i,j}^4 = \{x_j, e_{i,j}, a_{i,j}\},$$
$$c_{i,j}^5 = \{a_{i,j}, b_{i,j}, d_{i,j}, e_{i,j}\}.$$

- Let C = $\{c_i | 1 \leq \mathrm{i} \leq n\} \cup \{C_{\mathrm{i},j}^t | (v_i, v_j) \in \mathrm{E}, 1 \leq t \leq 5\}$,

$s = n + 4|E|$+k so that C and s can be constructed from G and k in polynomial time.

G has a vertex cover of size at most k iff C has a nomarl basis of cardinality at most s.

DFA→UFA is NP-complete

- To cover the set c_i, the basis B must contain either c_i and $\{x_i\}$,$\{y_i\}$
- $V_1 = \{v_i |\ both\ \{x_i\}\{y_i\} are\ in\ B\}$
- For fixed $\{x_i, y_i\} \in \mathrm{E}$ that at least four sets (in addition to sets c_i c_j $\{x_i\}$,$\{y_i\}$)are necessary to cover five sets $\{C^t_{\mathrm{i},j} | 1 \le t \le 5\}$, and four sets are sufficient iff at least one of v_i, v_j is in V_1.

Ruduce to our problem:

- Construct a DFA M as follow: The state of M is
  - $\Sigma = \{t | t \in c_i \text{ for } some\ i\} \cup = \{b_i | i = 1,2, \dots, n\}$.
  - $Q = \{q_0, q_1, \dots, q_n, q_f\}$;
  - $\delta: Q \times \Sigma \rightarrow \mathrm{Q}$ is defined as $\delta(q_0, b_i) = q_i, \delta(q_i, a^i_j) = q_f$ where 1≤i ≤n, 1 ≤ j ≤ n.
- Let k=s+2, claim that C has a normal basis of cardinality s iff L(M) is accepted by k-state UFA M'
  - $s_j \in \delta(q_0, b_i)$ iff r_j $belongs\ to\ c_i$;
  - $\delta'(s_j, a)$= q_f iff a ∈ r_j where $r_1, r_2, \dots, r_s$ be a normal basis of C.
- Conversely, assume M' is a minimal UFA.
  - Length=2, Q-$\{q'_0\}$ can be partitioned as $Q_1\ and\ Q_2$: $q'_0 \rightarrow Q_1 \rightarrow Q_2$
  - For each $q \in Q_1$, the set $B_q = \{a | q_f \in \delta(q, a)\}$ is a normal basis of C of size at most k-2.

Conclusion

- It's easy to show, by using 1 and 2, that problem is in NP. Let M and k be inputs. The nondeterministic algorithm will guess an NFA M' with at most k states. It will then check that M' is unambiguous by 2. Then by 1, it verifies they are equal and accepts.
- To prove NP-hardness, we reduce to the normal set basis problem which was proofed above.
- Thus, DFA$\rightarrow$UFA is in NP-Complete

References

Tao Jiang & B.Ravikumar, Minimal NFA problems are hard, 2-6