CS 311: Computational Structures

James Hook

October 2, 2014

2 NFA and DFA

2.1 Recall

- Motivation: A language to discuss computation.
- Deterministic Finite Automata (DFA). Given by cartoon or by formal mathematical structure. Finite number of states. Finite alphabet. State transition is a total function, with each state and symbol pair mapped to exactly one new state.
- Definitions of: acceptance, the language recognized by a DFA, the family of regular languages.
- Nondeterministic Finite Automata (NFA). Each state and symbol pair has a set of successor states (may be none, may be many).

2.2 Plan

- More NFA Examples
- Simulation of NFA by DFA
- Proof that NFAs represent exactly the Regular Languages.
- Discuss Problem Set 1.
- Regular languages closed under union.

2.3 NFA Examples

Ambiguous substring recognition.
Pattern followed by a pattern.
Lexical analysis motivated examples.

2.4 Simulation of NFA by DFA

Set up with a cartoon and coins.

Sets of states.

Powerset construction without ϵ transitions; proof sketch.

2.4.1 Proof of PowerSet construction

Sipser's construction is exemplary, but the proof asserts that "M obviously works correctly." That is a little on the informal side for my taste.

Construction: Given an NFA $N = (Q, \Sigma, \delta, q_0, F)$, construct DFA $M = (\mathcal{P}(Q), \Sigma, \delta', \{q_0\}, F')$ where $\delta'(R, a) = \bigcup_{r \in R} \delta(r, a)$ and $F' = \{R | R \cap F \neq \emptyset\}$.

Claim 2.1 $\mathcal{L}(N) = \mathcal{L}(M)$

To show this we establish we need to be able to show (1) $w \in \mathcal{L}(N) \Rightarrow \mathcal{L}(M)$ and (2) $w \in \mathcal{L}(M) \Rightarrow \mathcal{L}(N)$.

To show (1), we need to show that when there is a sequence of NFA states $r_1, r_2, \ldots, r_n$ that witness acceptance of w that by N there is a corresponding sequence of DFA states $R_1, R_2, \ldots, R_n$ that witnesses acceptancy of w by M. To prove that, we need to establish by induction that $r_i \in R_i$ at every step.

Claim 2.2 *If $r_0, r_1, \ldots, r_k$ is a sequence of NFA states satisfying conditions (1) and (2) of the definition of acceptance, then there is a corresponding sequence of DFA states $R_0, R_1, \ldots, R_k$ also satisfying conditions (1) and (2) such that for all $i \leq k$, $r_i \in R_i$.*

Proof by induction on k.

Basis. When $k = 0$ by property (1) we know $r_1 = q_0$. By construction, $R_0 = \{q_0\}$. Clearly $q_0 \in \{q_0\}$.

Step. Assume for k to show for $k + 1$. Assume $w = w'a$ where a is a symbol in Σ. Given $r_0, r_1, \ldots, r_k, r_{k+1}$ satisfying (1) and (2) we get $R_0, R_1, \ldots, R_k$ by induction and $r_{k+1} \in \delta(r_k, a)$ from (2). By construction, $R_{k+1} = \delta'(R_k, a) = \bigcup_{r \in R_k} \delta(r, a)$. Since $r_k \in R_k$, we have $\delta(r_k, a) \subseteq R_{k+1}$. Hence $r_{k+1} \in R_{k+1}$ as required.

Corollary 2.3 *If N accepts w then M accepts w.*

If N accepts w then by definition of acceptance there is a sequence satisfying conditions (1), (2), and (3). By Claim 2.2 there is a corresponding sequence for M with $r_n \in R_n$. Since $r_n \in F$, $R_n \cap F \neq \emptyset$, hence $R_n \in F'$ and thus M accepts w.

To show (2) we need a similar argument. This time we show that if the DFA can get from R_0 to R_k then the NFA can get to any state s in R_k by a path of the form $r_0, r_1, \ldots, r_k$.

Claim 2.4 *If $R_0, R_1, \ldots, R_k$ is a sequence of DFA states satisfying conditions (1) and (2) then for every $s \in R_k$ there is a sequence $r_0, r_1, \ldots, r_k$ where $s = r_k$ and the sequence satisfies conditions (1) and (2).*

Proof by induction on k.

Basis. When $k = 0$ then $r_1 = q_0$ and $R_1 = \{q_0\}$ as before.

Step. Assume for k to show for $k + 1$. Let w be of the form $w'a$. Given the sequence $R_0, R_1, \ldots, R_k, R_{k+1}$ we get by induction that all $s' \in R_k$ are reachable by a sequence of states of N. Consider an arbitrary $s \in R_{k+1}$. By construction $R_{k+1} = \delta'(R_k, a) = \bigcup_{r \in R_k} \delta(r, a)$. Hence any $s \in R_{k+1}$ must be reached from some NFA state $s' \in R_k$. Let $r_0, r_1, \ldots, r_k$ be the sequence reaching s', extend that with $r_{k+1} = s$ to complete the sequence reaching s. This completes the proof of the claim.

Corollary 2.5 *If M accepts w then N accepts w.*

If M accepts w then by definition of acceptance there is a sequence satisfying conditions (1), (2), and (3). Let s be an element of both R_n and F. By Claim 2.4 there is a corresponding sequence $r_0, r_1, \ldots, r_n$ where $r_n = s$ satisfying conditions (1) and (2). Since $s \in F$ this satisfies (3). Hence N accepts w as required.

2.4.2 ϵ-closure

Note discussion in text.