Informatics 1 - Computation & Logic: Tutorial 8

Computation: Subset Procedure

Week 10: 18-22 November 2013

Please attempt the entire worksheet in advance of the tutorial, and bring with you all work, including (if a computer is involved) printouts of code and test results. Tutorials cannot function properly unless you do the work in advance.

You may work with others, but you must understand the work; you can't phone a friend during the exam.

Assessment is formative, meaning that marks from coursework do not contribute to the final mark. But coursework is not optional. If you do not do the coursework you are unlikely to pass the exams.

Attendance at tutorials is **obligatory**; please let your tutor know if you cannot attend.

The Subset Procedure

Recall the following definition:

A finite state machine is an ordered 5-tuple $\langle Q, \Sigma, s_0, F, \delta \rangle$*, where:*

- Q *is a finite set of states;*
- Σ *is an alphabet of input symbols;*
- $s_0 \in Q$ *is the initial state;*
- $F \subseteq Q$ *is the set of accept states; and*
- δ *is a set of transitions, essentially a subset of* $Q \times \Sigma \times Q$.

A finite state machine $\langle Q, \Sigma, s_0, F, \delta\rangle$ *is deterministic if and only if for every* $q \in Q$ *and every* $\sigma \in \Sigma$*, there is just one transition* $\langle q, \sigma, q'\rangle \in \delta$*.*

Every non-deterministic FSM can be converted into a deterministic FSM that accepts exactly the same set of strings, by means of an algorithm known as the *subset procedure.*

The formal definition states that the result of applying the subset procedure to FSM $\langle Q, \Sigma, s_o, F, \delta\rangle$ is the deterministic FSM $\langle Q', \Sigma, \{s_o\}, F', \delta'\rangle$, where:

- Q' is a set of subsets of Q i.e. superstates
- F' is the set of all and only superstates in Q' which contain at least one state in F
- δ' is the set of derived transitions — given the original FSM $\langle Q, \Sigma, s_o, F, \delta\rangle$, in the new, derived machine, there is a transition from superstate A by means of symbol $\sigma \in \Sigma$ to the following superstate:

  $$\{q \mid \text{for some } q' \in A, \langle q', \sigma, q\rangle \in \delta\}$$

  i.e. the set of all states in the original machine that can be reached from *some* state in A by means of a transition labelled σ.

Take for example the following FSM, which accepts all and only the strings over $\{b, c\}$ that begin with b and end with c:

$$\mathcal{M}_1 = \langle\{0, 1, 2\}, \{b, c\}, 0, \{2\}, \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\rangle$$

It may help to draw out this FSM as a diagram. This machine is non-deterministic because there are two distinct transitions from state 1 for the symbol c.

Let us now consider how to convert $\mathcal{M}_1$ into a deterministic FSM, $\mathcal{M}_2$, that accepts exactly the same set of strings.

First of all, note that the set of states in $\mathcal{M}_1$ is $\{0, 1, 2\}$. The states in the deterministic machine $\mathcal{M}_2$ will be *subsets* of those states, which we call *superstates.*

Step 1: Identify initial superstate

The first step in converting $\mathcal{M}_1$ into a deterministic machine is to identify the initial superstate of the new machine. To do this we just take the initial state of $\mathcal{M}_1$, i.e. 0, and convert it into a superstate containing only this state — $\{0\}$. Thus, the new machine, $\mathcal{M}_2$, starts off as follows:

$$\mathcal{M}_2 = \langle\{\{0\}\}, \{b, c\}, \{0\}, \emptyset, \emptyset\rangle$$

Again, it might be useful to draw this partial machine. Note that as of yet the machine $\mathcal{M}_2$ does not contain any accept states or transitions. It has just the single state, $\{0\}$, that is also the initial state.

Step 2: Incrementally add transitions

The next part of the subset procedure involves incrementally building up the partial FSM until every superstate has exactly one transition for each symbol in the alphabet $\{b, c\}$. First, we identify those transitions that are missing — the partial machine has one state $\{0\}$ and this state lacks two transitions: one for symbol b and one for symbol c. Thus we must find:

- a transition from superstate $\{0\}$ for symbol b, and,
- a transition from superstate $\{0\}$ for symbol c

Given some non-deterministic FSM $\langle Q, \Sigma, s_0, F, \delta\rangle$, a superstate $A \in \wp(Q)$ and a symbol $\sigma \in \Sigma$, the superstate in the new machine that is reached from superstate A by means of symbol σ is defined as follows:

$$\{q \mid \text{for some } q' \in A, \langle q', \sigma, q\rangle \in \delta\}$$

Thus:

- the transition from superstate $\{0\}$ for symbol b is:

  $$\{q \mid \text{for some } q' \in \{0\}, \langle q', b, q\rangle \in \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\} = \{1\}$$

- the transition from superstate $\{0\}$ for symbol c is:

  $$\{q \mid \text{for some } q' \in \{0\}, \langle q', c, q\rangle \in \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\} = \emptyset$$

Thus, we have two new transitions to add to our partial machine: $\langle\{0\}, b, \{1\}\rangle$ and $\langle\{0\}, c, \emptyset\rangle$. Adding the new transitions and the new superstates $\{1\}$ and $\emptyset$ to $\mathcal{M}_2$, the partial machine is now:

$$\mathcal{M}_2 = \langle\{\{0\}, \{1\}, \emptyset\},\ \{b, c\},\ \{0\},\ \emptyset,\ \{\langle\{0\}, b, \{1\}\rangle,\ \langle\{0\}, c, \emptyset\rangle\}\rangle$$

This partial machine still has transitions missing, so we continue the procedure:

- the transition from superstate $\{1\}$ for symbol b is:

  $$\{q \mid \text{for some } q' \in \{1\}, \langle q', b, q\rangle \in \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\} = \{1\}$$

- the transition from superstate $\{1\}$ for symbol c is:

  $$\{q \mid \text{for some } q' \in \{1\}, \langle q', c, q\rangle \in \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\} = \{1, 2\}$$

- the transition from superstate $\emptyset$ for symbol b is:

  $$\{q \mid \text{for some } q' \in \emptyset, \langle q', b, q\rangle \in \{\langle 0, b, 1\rangle, \langle 1, b, 1\rangle, \langle 1, c, 1\rangle, \langle 1, c, 2\rangle\}\} = \emptyset$$

- the transition from superstate $\emptyset$ for symbol c is:

$$\{q \mid \text{for some } q' \in \emptyset, \langle q', c, q\rangle \in \{\langle 0,b,1\rangle, \langle 1,b,1\rangle, \langle 1,c,1\rangle, \langle 1,c,2\rangle\}\} = \emptyset$$

Thus the superstate $\{1,2\}$ is added to machine $\mathcal{M}_2$, along with four new transitions: $\langle\{1\}, b, \{1\}\rangle$, $\langle\{1\}, c, \{1,2\}\rangle$, $\langle\emptyset, b, \emptyset\rangle$, and $\langle\emptyset, c, \emptyset\rangle$. The partial machine is now (again, it may be useful to draw out the partial machine):

$$\mathcal{M}_2 = \langle\{\{0\}, \{1\}, \emptyset, \{1,2\}\},\ \{b,c\},\ \{0\},\ \emptyset,\ \delta\rangle$$

where $\delta = \{\langle\{0\}, b, \{1\}\rangle,\ \langle\{0\}, c, \emptyset\rangle,\ \langle\{1\}, b, \{1\}\rangle,\ \langle\{1\}, c, \{1,2\}\rangle,\ \langle\emptyset, b, \emptyset\rangle,\ \langle\emptyset, c, \emptyset\rangle\}$

There is *still* one superstate $\{1,2\}$ which lacks the appropriate transitions, so we continue:

- the transition from superstate $\{1,2\}$ for symbol b is:

$$\{q \mid \text{for some } q' \in \{1,2\}, \langle q', b, q\rangle \in \{\langle 0,b,1\rangle, \langle 1,b,1\rangle, \langle 1,c,1\rangle, \langle 1,c,2\rangle\}\} = \{1\}$$

- the transition from superstate $\{1,2\}$ for symbol c is:

$$\{q \mid \text{for some } q' \in \{1,2\}, \langle q', c, q\rangle \in \{\langle 0,b,1\rangle, \langle 1,b,1\rangle, \langle 1,c,1\rangle, \langle 1,c,2\rangle\}\} = \{1,2\}$$

So we add a further two transitions to $\mathcal{M}_2$: $\langle\{1,2\}, b, \{1\}\rangle$ and $\langle\{1,2\}, c, \{1,2\}\rangle$. The machine is now:

$$\mathcal{M}_2 = \langle\{\{0\}, \{1\}, \emptyset, \{1,2\}\},\ \{b,c\},\ \{0\},\ \emptyset,\ \delta_2\rangle$$

where

$$\begin{aligned}\delta_2 = \ & \{\langle\{0\}, b, \{1\}\rangle,\ \langle\{0\}, c, \emptyset\rangle,\ \langle\{1\}, b, \{1\}\rangle,\ \langle\{1\}, c, \{1,2\}\rangle,\\ & \langle\emptyset, b, \emptyset\rangle,\ \langle\emptyset, c, \emptyset\rangle,\ \langle\{1,2\}, b, \{1\}\rangle,\ \langle\{1,2\}, c, \{1,2\}\rangle\}\end{aligned}$$

The transition set of the new machine is now complete, since every superstate has exactly one transition for each symbol in the alphabet $\{b,c\}$.

Step 3: Identify accept superstates

The only thing that remains to do is to identify the set of accept states in the new, deterministic FSM. Basically, *all* superstates in the new FSM that contains at least *one* of the accept states in the original non-deterministic FSM are accept states of the new, derived machine. The set of accept states in $\mathcal{M}_1$ is $\{2\}$, so the only accept state in $\mathcal{M}_2$ is the superstate $\{1,2\}$.

Step 4: Full definition of deterministic FSM

In conclusion, the result of applying the subset procedure to the non-deterministic FSM $\mathcal{M}_1$ is the following *deterministic* machine $\mathcal{M}_2$:

$$\mathcal{M}_2 = \langle \{\{0\}, \{1\}, \emptyset, \{1,2\}\},\ \{b, c\},\ \{0\},\ \{\{1,2\}\},\ \delta_2 \rangle$$

where

$$\begin{aligned} \delta_2 = \ & \{\langle \{0\}, b, \{1\}\rangle,\ \langle \{0\}, c, \emptyset\rangle,\ \langle \{1\}, b, \{1\}\rangle,\ \langle \{1\}, c, \{1,2\}\rangle, \\ & \langle \emptyset, b, \emptyset\rangle,\ \langle \emptyset, c, \emptyset\rangle,\ \langle \{1,2\}, b, \{1\}\rangle,\ \langle \{1,2\}, c, \{1,2\}\rangle\} \end{aligned}$$

1. Using the subset procedure, convert the following non-deterministic FSM:

   $$\mathcal{M}_1 = \langle \{1,2\},\ \{a, b\},\ 1,\ \{2\},\ \{\langle 1, a, 1\rangle, \langle 1, b, 1\rangle, \langle 1, b, 2\rangle\}\rangle$$

   to a deterministic FSM, $\mathcal{M}_2$, that accepts the same language.

   (a) Draw the $\mathcal{M}_1$ machine:

   (b) Give the complete set of states, Q_2, of $\mathcal{M}_2$:

   (c) What is the start state, s_0, of $\mathcal{M}_2$?

   (d) Give the set of transitions, δ_2, of $\mathcal{M}_2$:

(e) What is the set of accepting states, F_2, of $\mathcal{M}_2$?

(f) Now give the full definition of $\mathcal{M}_2$:

(g) Finally, draw the $\mathcal{M}_2$ machine:

2. Convert the following non-deterministic FSM:

$$\mathcal{M}_3 = \langle\{1,2,3,4\},\{a,b\},\ 1,\ \{4\},\delta_3\rangle$$

$\delta_3 = \{\langle 1,a,1\rangle,\langle 1,b,1\rangle,\langle 1,b,2\rangle,\langle 1,a,3\rangle,\langle 2,b,4\rangle,\langle 3,a,4\rangle,\langle 4,a,4\rangle,\langle 4,b,4\rangle\}\rangle$ to a deterministic FSM, $\mathcal{M}_4$, that accepts the same language.

(a) Draw the $\mathcal{M}_3$ machine:

(b) What is the start state, s_0, of $\mathcal{M}_4$?

(c) Give the set of transitions, δ_4, of $\mathcal{M}_4$:

(d) Give the set of states, Q_4, of $\mathcal{M}_4$:

(e) What is the set of accepting states, F_4, of $\mathcal{M}_4$?

(f) Now give the full definition of $\mathcal{M}_4$:

(g) Finally, draw the $\mathcal{M}_4$ machine:

This tutorial exercise sheet was written by Mark McConville, revised by Paolo Besana, and extended by Thomas French and Areti Manataki. Send comments to `A.Manataki@ed.ac.uk`