I.4.1. String Recognition (13.1)

I.4.2. Nondeterministic Finite State Automata (13.2)

I.4.3. Examples of Automata (13.3)

I.4.4. Automata with Empty Move Transitions (13.4)

I.4.5. Deterministic Finite State Automata (13.6)

Example Recognising Strings, Step 1

We want to define a program which recognises strings “start” and “stop”. We start by defining a program which recognises the letter “s”. This can be given by a system given by the following diagram, which will be our first automaton (automata will be introduced soon).

This automaton has the following ingredients.

- States $q_0$, $q_1$.
- State $q_0$ is the starting state, indicated by the arrow into it coming from nowhere.
- State $q_1$, the state indicating that we have recognised letter “s”.
- A transition, which when recognising letter “s”, goes from state $q_0$ to $q_1$. 

Anton Setzer
(Based on a book draft by J. V. Tucker and K. Stephenson)
Dept. of Computer Science, Swansea University

http://www.cs.swan.ac.uk/~csetzer/lectures/languageComputation/10/index.html

December 8, 2010
Example Recognising Strings, Step 2

In order to recognise the letter "t", we extend this automaton as follows:

- $q_0$ is the start state.
- $q_1$ indicates we have read string "s".
- $q_2$ indicates we have read string "st".

Example Recognising Strings, Step 3

For the 3rd letter, we have two choices: "a" as part of the word "start", and "o" as part of the word "stop".

- $q_0$ is the start state.
- $q_1$ indicates we have read string "s".
- $q_2$ indicates we have read string "st".
- $q_3$ indicates we have read string "sta".
- $q_6$ indicates we have read string "sto".
We can complete our automaton and obtain the following:

This diagram contains a new ingredient: States $q_5$ and $q_7$ are \textbf{accepting states}. If we have processed a word and reached such a state then the word is accepted as a string of the language of the automaton.
(In our example $L = \{\text{start, stop}\}$.)

\textbf{Automata with Loops}

In order to recognise infinite languages, we need automata with loops. The following automaton recognises $L = \cdots$.

The following automaton recognises $L = \cdots$.
Automata with Loops

The following automaton recognises $L = ????$

- The automaton chooses in state $q_0$, non-deterministically, when in state $q_0$ and recognising a letter $s$, whether to go to $q_1$ or $q_6$.
- The accepted language is the set of strings such that for each of them we obtain an accepting state for at least one non-deterministic choice.

Nondeterministic Automata

The language \{start, stop\} can be as well recognised by the following nondeterministic automaton:

- If we try to accept the word “stop” by moving $q_0 \xrightarrow{s} q_1 \xrightarrow{t} q_2$ we get stuck at $q_2$.
- That we fail to accept a word for one specific non-deterministic choice doesn’t imply that this word is not in the language.
- A word is not accepted only if for all non-deterministic choices the corresponding run of the automaton doesn’t accept the string.
Why Nondeterministic Automata?

When translating regular grammars into automata, we will obtain non-deterministic automata.

We will show later that from a non-deterministic automaton we can obtain an equivalent deterministic automaton.

Non-deterministic machine models play an important role in the theory of algorithms and complexity.

In some cases (as for automata), deterministic and non-deterministic are equivalent.

Sometimes they are not.

In other cases it is an open problem whether they are equivalent.

---

**Definition NFA**

A non-deterministic finite state automaton, in short NFA $(Q, q_0, F, T, \rightarrow)$ is given by

- A finite set $Q$ of states.
- A single initial state $q_0$.
- A set $F \subseteq Q$ of accepting states.
- A finite set of terminal symbols $T$.
- $\rightarrow$ is a relation between states $q \in Q$, $a \in T$ and states $q' \in Q$. We write $q \xrightarrow{a} q'$ if $q, a, q'$ are in this relation.

More formally $\rightarrow$ is a subset of $Q \times T \times Q$ and $q \xrightarrow{a} q'$ is an abbreviation for $(q, a, q') \in \rightarrow$.

---

**Presentation of NFA by Picture**

An NFA can be presented by a picture, like the following diagram:

- The arrow from nowhere into state $q_0$ denotes that $q_0$ is the initial state.
- Circles denote states.
- Arrows from a state $q$ to $q'$ labelled by $a$ mean that $q \xrightarrow{a} q'$.
- Double circle like for $q_5$ and $q_9$ denote the accepting states.
- The alphabet needs to be stated, unless all elements of the alphabet occur as labels of arrows, or it is clear from the context.

Alphabet $= \{a, \ldots, z\}$
Presented by Table

automaton $A$

states $q_0, \ldots, q_n$

terminals $a_0, \ldots, a_m$

start $q_i$

final $q_{i_0}, q_{i_1}, \ldots, q_{i_m}$

transitions $q_{k_1} \xrightarrow{a_{i_1}} q_{m_1}$

Example

The automaton for alphabet $\{a, \ldots, z\}$ given by the diagram

is represented by the table on the next slide.

The Extended Transition Relation

We define an extended transition relation which determines, whether the NFA from a state $q$ when reading word $w \in T^*$ can reach state $q'$:

Definition

Let $A = (Q, q_0, F, T, \rightarrow)$ be an NFA. Then we define $\rightarrow^*$ between states $q \in Q$, words $w \in T^*$ and $q' \in Q$.

We write $q \xrightarrow{w}^* q'$, if $q, w, q'$ are in this relation.

- If $w = a_0 \cdots a_n$, then

$$q \xrightarrow{w}^* q' \text{ iff } q \xrightarrow{a_0} q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} q_n = q'$$

for some $q_i \in Q$

Especially $q \xrightarrow{\epsilon} q'$ iff $q' = q$. 
The Extended Transition Relation

Again, more formally $\rightarrow^*$ is a subset of $Q \times T^* \times Q$ and $q \xrightarrow{w^*} q'$ means that $(q, w, q') \in \rightarrow^*$

Example

For each of the words given in the following all possible transitions from $q_0$ will be given:

- $q_0 \xrightarrow{} q_0$,
- $q_0 \xrightarrow{s} q_1$, $q_0 \xrightarrow{s} q_6$,
- $q_0 \xrightarrow{st} q_2$, $q_0 \xrightarrow{st} q_7$,
- $q_0 \xrightarrow{sta} q_3$,
- $q_0 \xrightarrow{star} q_4$,
- $q_0 \xrightarrow{start} q_5$,
- $q_0 \xrightarrow{sto} q_8$,
- $q_0 \xrightarrow{stog} q_9$.

The Language accepted by an NFA

Definition

Let $A = (Q, q_0, F, T, \rightarrow)$ be an NFA. The language accepted by $A$ is defined as

$$L(A) := \{ w \in T^* \mid q_0 \xrightarrow{w^*} q \text{ for some } q \in F \}$$
I.4.2. Nondeterministic Finite State Automata (13.2)

Operational Understanding of NFAs

To an NFA $A = (Q, q_0, F, T, \rightarrow)$ corresponds a non-deterministic program for checking whether an input string is in $L(A)$. The program will have

- a variable $q \in Q$,
- and a pointer to a position in the input string,
- a Boolean variable $stopped$, which will be true if the automaton has stopped without having reached the end of the input string.

A run of the program on an input string $s$ is as follows:

- Initially
  - $q := q_0$ (the initial state),
  - $p$ points to the beginning of $s$,
  - $stopped := false$.

- While $p$ doesn’t point to the end of $s$ and $stopped = false$ do:
  - Let $a$ be the next symbol from the input string.
  - If there is no $q' \text{ s.t. } q \xrightarrow{a} q'$,
    - $stopped := true$.
  - Otherwise choose non-deterministically an element $q'$ s.t. $q \xrightarrow{a} q'$, and then define
    - $q := q'$
    - $p$ is set to the next symbol of the input string,
    - $stopped := false$.

- End of the while loop.

If $stopped = true$ (therefore we have reached the end of the string), the string is not accepted in this run.

Otherwise the string is accepted if $q \in F$.

A string is accepted if there exist a sequence of non-deterministic choices, such that the string is accepted in the corresponding run.

One can now see that a string $s$ is accepted by some run of the program for automaton $A$ iff $s \in L(A)$. 
I.4.3. Examples of Automata (13.3)

Example 1: An Automaton accepting \{1, 2, 3\}

Display style:

- **automaton**: \( A^{1,2,3} \)
- **states**: \( q_0, q_1, q_2, q_3 \)
- **terminals**: 1, 2, 3
- **start**: \( q_0 \)
- **final**: \( q_1, q_2, q_3 \)
- **transitions**:
  - \( q_0 \xrightarrow{1} q_1 \)
  - \( q_0 \xrightarrow{2} q_2 \)
  - \( q_0 \xrightarrow{3} q_3 \)

Simplified Version
Example 2: Automaton accepting ???
Example 3: Automaton accepting ???

Display Style

automaton $A$

states $q_0, q_1, q_2$

terminals $0, \ldots, 9$

start $q_0$

final $q_1, q_2$

transitions

$q_0 \xrightarrow{0} q_2$
$q_0 \xrightarrow{a} q_1, (a \in \{1, \ldots, 9\})$
$q_1 \xrightarrow{a} q_1, (a \in \{0, \ldots, 9\})$

Material for this Section Moved to Additional Material

The material for this section has been moved to Additional Material
I.4.5. Deterministic Finite State Automata (13.6)

**Definition DFA**

Let $A = (Q, q_0, F, T, \rightarrow)$ be an NFA. $A$ is a **deterministic finite state automaton**, in short DFA, if for all $q \in Q$, there exist at most one $q'$ s.t.

$$a \rightarrow q'$$

So deterministic finite state automata are those automata corresponding to real programs: we have never to make a choice.

**Example (DFA)**

The following NFA is a DFA:

![Diagram of a DFA]

Note that there is no transition from $q_0$ labelled by letter $\neq t$.

For a DFA from each state and element of the alphabet there need to be at most one transition – it is possible to have no transition.

**Example (not a DFA)**

The following NFA is not a DFA:

![Diagram of an NFA that is not a DFA]
Theorem I.4.5.1.

Let $A$ be an NFA. Then there exists a DFA $A'$ s.t. $L(A) = L(A')$. $A'$ can be computed from $A$.

Example

Consider the NFA above which was not an DFA:

Definition of the DFA

- We define a DFA as follows:
  - The states are all possible sets of states.
    - Note that since the set of states is finite, the set of sets of states is finite, too.
  - The initial state is the set of initial states, here $\{q_0\}$.

- From a set of states $B$ we can make an $a$-transition to the set of states $q$, which can be reached from any state $q \in B$ by an $a$-transition.
  - For instance
    - $\{q_0\} \xrightarrow{a} \{q_1, q_6\}$ since the set of states we can reach by an $a$-transition from $q_0$ are states $q_1, q_6$.
    - $\{q_1, q_6\} \xrightarrow{t} \{q_2, q_7\}$ since the set of states we can reach by a $t$-transition from any of the states $q_1$ or $q_6$ are the states $q_2, q_7$.
  - A set of states is an accepting state, if at least of the states in it is in $F$.
    - For instance $\{q_9\}$ or $\{q_5\}$ or $\{q_4, q_9\}$ are accepting.
Definition of the DFA

- ∅ is not an accepting state, and ∅ → ∅.
  - So ∅ is a sink of the DFA. It’s non-accepting and we can’t escape from it.
  - We can omit it and all transitions into it.
- We can as well omit all states which are unreachable from the initial state.
- So we start from the initial state \{q_0\}, and successively determine all transitions of the DFA which don’t end up in ∅.

Corresponding DFA

We obtain the following automaton:

Note that this was the original DFA above (up to renaming of states).

Example 2

Consider the following NFA accepting \(L = \ldots\):

Corresponding DFA

Consider the following NFA accepting \(L = \ldots\):
Proof Idea

- Now one can see that in the DFA we have that

\[ \{q_0\} \xrightarrow{w,*} B \]

where

\[ B = \{ q \mid q_0 \xrightarrow{w,*} q \} \]

so \( B \) is the set of states which we can reach from \( B \) by a \( w \)-transition.

- Therefore in the DFA we reach an accepting state (one containing an element of \( F \)) by a \( w \)-transition, iff the NFA can reach an accepting state by a \( w \)-transition.

- Therefore the language accepted by this DFA is the same as the language accepted by the original NFA.

- A formal proof can be found in the Additional Material.