Finite Automata Theory and Formal Languages

TMV027/DIT321– LP4 2018

Lecture 4

Ana Bove

March 23rd 2018

Recap: Formal Proofs

- How formal should a proof be? Depends on its purpose...
- ... but should be convincing ...
  ... and the validity of each step should be easily understood;
- One proves the conclusions assuming the validity of the hypotheses!
- Different kind of proofs (contradiction, contrapositive, counterexample, induction, ...);
- Simple/strong induction allows to prove a property over *all* Natural numbers;
- Sometimes we prove several properties that depend on each other (mutual induction);
- Inductive definitions generate possibly infinite sets with finite elements: Booleans, Natural numbers, lists, trees, ...
- We can recursively defined functions over inductive sets.

Overview of Today's Lecture

- Structural induction;
- Concepts of automata theory.

Contributes to the following learning outcome:

- Prove properties of languages, grammars and automata with rigorously formal mathematical methods.

Inductively Defined Sets (Recap)

To define a set S by induction we need to specify:

Base cases: $\mathsf{e}_1[\ldots], \ldots, \mathsf{e}_\mathsf{m}[\ldots] \in S$;

Inductive steps: Given $\ldots$ and $s_1, \ldots, s_{k_i} \in S$,
then $\mathsf{c}_1[\ldots, s_1, \ldots, s_{k_1}], \ldots, \mathsf{c}_\mathsf{n}[\ldots, s_1, \ldots, s_{k_n}] \in S$;

Closure: There is no other way to construct elements in S.
(We will usually omit this part.)

Alternatively,

$$\frac{\ldots}{\mathsf{e}_1[\ldots] \in S} \quad \ldots \quad \frac{\ldots}{\mathsf{e}_\mathsf{m}[\ldots] \in S}$$

$$\frac{\cdots \quad s_1, \ldots, s_{k_1} \in S}{\mathsf{c}_1[\ldots, s_1, \ldots, s_{k_1}] \in S} \quad \ldots \quad \frac{\cdots \quad s_1, \ldots, s_{k_n} \in S}{\mathsf{c}_\mathsf{n}[\ldots, s_1, \ldots, s_{k_n}] \in S}$$

Proofs by Structural Induction

Generalisation of mathematical induction to other inductively defined sets such as lists, trees, ...

VERY useful in computer science: it allows to prove properties over the (finite) elements in a data type!

Given an inductively defined set S, to prove $\forall\, s \in S.\, P(s)$ then:

Base cases: We prove that $P(\mathrm{e}_1[\ldots]), \ldots, P(\mathrm{e}_\mathrm{m}[\ldots])$;

Inductive steps: Assuming $P(s_1), \ldots, P(s_{k_i})$ (our *inductive hypotheses* IH),
we prove $P(\mathrm{c}_1[\ldots, s_1, \ldots, s_{k_1}]), \ldots, P(\mathrm{c}_\mathrm{n}[\ldots, s_1, \ldots, s_{k_n}])$;

Closure: $\forall\, s \in S.\, P(s)$.
(We will usually omit this part.)

Inductive Sets and Structural Induction

Inductive definition of S:

$$\frac{\cdots}{\mathrm{e}_1[\ldots] \in S} \quad \cdots \quad \frac{\cdots}{\mathrm{e}_\mathrm{m}[\ldots] \in S} \quad \frac{\cdots \;\; s_1, \ldots, s_{k_1} \in S}{\mathrm{c}_1[\ldots, s_1, \ldots, s_{k_1}] \in S} \quad \cdots \quad \frac{\cdots \;\; s_1, \ldots, s_{k_n} \in S}{\mathrm{c}_\mathrm{n}[\ldots, s_1, \ldots, s_{k_n}] \in S}$$

Inductive principle associated to S:

$$\dfrac{\begin{array}{rl} \text{base cases} & \left\{ \begin{array}{l} P(\mathrm{e}_1[\ldots]) \\ \quad\vdots \\ P(\mathrm{e}_\mathrm{m}[\ldots]) \end{array} \right. \\ \text{inductive steps} & \left\{ \begin{array}{l} \forall\, \ldots, s_1, \ldots, s_{k_1} \in S.\, \overbrace{P(s_1), \cdots, P(s_{k_1})}^{\text{IH}} \Rightarrow P(\mathrm{c}_1[\ldots, s_1, \ldots, s_{k_1}]) \\ \qquad\qquad\qquad\vdots \\ \forall\, \ldots, s_1, \ldots, s_{k_n} \in S.\, P(s_1), \cdots, P(s_{k_n}) \Rightarrow P(\mathrm{c}_\mathrm{n}[\ldots, s_1, \ldots, s_{k_n}]) \end{array} \right. \end{array}}{\forall\, s \in S.\, P(s)}$$

Example: Structural Induction over Lists

We can now use recursion to define functions over an inductively defined set and then prove properties about these functions by structural induction.

Recall recursive definitions of append and length functions over lists:

$$\begin{aligned} [] \mathbin{++} ys &= ys & \qquad \text{len } [] &= 0 \\ (x : xs) \mathbin{++} ys &= x : (xs \mathbin{++} ys) & \qquad \text{len } (x : xs) &= 1 + \text{len } xs \end{aligned}$$

Proposition: $\forall\, xs, ys \in \text{List } A.\, \text{len } (xs \mathbin{++} ys) = \text{len } xs \mathbin{++} \text{len } ys.$

Proof: By structural induction on xs.
$P(xs)$ is $\forall\, ys \in \text{List } A.\, \text{len } (xs \mathbin{++} ys) = \text{len } xs \mathbin{++} \text{len } ys$.

Base case: We prove $P[]$.
Inductive step: We show $\forall\, x \in A. \forall\, xs \in \text{List } A.\, P(xs) \Rightarrow P(x : xs)$.

Closure: $\forall\, xs \in \text{List } A.\, P(xs)$.

Example: Structural Induction over Lists

Proposition: $\forall\, xs \in \text{List } A.\, xs \mathbin{++} [] = xs.$

Proof: By structural induction on xs.
$P(xs)$ is $xs \mathbin{++} [] = xs$.

Base case: We prove $P[]$.
Inductive step: We show $\forall\, x \in A. \forall\, xs \in \text{List } A.\, P(xs) \Rightarrow P(x : xs)$.
Closure: $\forall\, xs \in \text{List } A.\, P(xs)$.

Proposition: $\forall\, xs, ys, zs \in \text{List } A.\, (xs \mathbin{++} ys) \mathbin{++} zs = xs \mathbin{++} (ys \mathbin{++} zs).$

Proof: By structural induction on xs.
$P(xs)$ is $\forall\, ys, zs \in \text{List } A.\, (xs \mathbin{++} ys) \mathbin{++} zs = xs \mathbin{++} (ys \mathbin{++} zs)$.

Base case: We prove $P[]$.
Inductive step: We show $\forall\, x \in A. \forall\, xs \in \text{List } A.\, P(xs) \Rightarrow P(x : xs)$.
Closure: $\forall\, xs \in \text{List } A.\, P(xs)$.

Example: Structural Induction over Trees

Let us (recursively) define functions counting the number of edges and of nodes of a tree:

$$\begin{aligned} \mathrm{ne}(x) &= 0 \\ \mathrm{ne}(x, t_1, \ldots, t_k) &= k + \\ &\quad \mathrm{ne}(t_1) + \ldots + \mathrm{ne}(t_k) \end{aligned} \qquad \begin{aligned} \mathrm{nn}(x) &= 1 \\ \mathrm{nn}(x, t_1, \ldots, t_k) &= 1 + \\ &\quad \mathrm{nn}(t_1) + \ldots + \mathrm{nn}(t_k) \end{aligned}$$

Proposition: $\forall\, t \in \text{Tree } A.\, \mathrm{nn}(t) = 1 + \mathrm{ne}(t)$.

Proof: By structural induction on t.
$P(t)$ is $\mathrm{nn}(t) = 1 + \mathrm{ne}(t)$.

Base case: We prove $P(x)$.
Inductive step: We show
$\forall\, x \in A. \forall\, t_1, \ldots, t_k \in \text{Tree } A.\, P(t_1), \ldots, P(t_k) \Rightarrow P(x, t_1, \ldots, t_k)$.
Closure: $\forall\, t \in \text{Tree } A.\, P(t)$.

Proofs by Induction: Overview of the Steps to Follow

1. State property P to prove by induction.
   Might be more general than the actual statement we need to prove!
2. **Determine and state the method to use in the proof!!!!**
   Example: Mathematical induction on the length of the list, course-of-values induction on the height of a tree, structural induction over a certain element, ...
3. Identify and state base case(s).
   Could be more than one! Not always trivial to determine.
4. Prove base case(s).
5. Identify and state IH!
   Will depend on the method to be used (see point 2).
6. Prove inductive step(s).
7. (State closure.)
8. Deduce your statement from P (if needed).

Central Concepts of Automata Theory: Alphabets

Definition: An *alphabet* is a finite, non-empty set of symbols, usually denoted by Σ.

The number of symbols in Σ is denoted as $|\Sigma|$.

Notation: We will use $a, b, c, \ldots$ to denote symbols.

Note: Alphabets will represent the observable events of an automaton.

Example: Some alphabets:

- on/off-switch: $\Sigma = \{\text{Push}\}$;
- parity counter: $\Sigma = \{p_0, p_1\}$;
- complex vending machine: $\Sigma = \{5\ kr, 10\ kr, \text{tea}, \text{coffee}\}$.

Strings or Words

Definition: *Strings/Words* are finite sequence of symbols from some alphabet.

Notation: We will use $w, x, y, z, \ldots$ to denote words.

Note: Words will represent the *behaviour* of an automaton.

Example: Some behaviours:

- on/off-switch: Push Push Push Push;
- parity counter: p_0p_1 or $p_0p_0p_0p_1p_1p_0$;
- comlex vending machine: 5 kr 5 kr coffee 10 kr coffee 5 kr tea.

Note: Some words do NOT represent *behaviour* though ...

Example: complex vending machine: tea 5 kr coffee.

Inductive Definition of Σ^*

Definition: Σ^* is the set of *all words* for a given alphabet Σ.

This can be described inductively in the following two ways:

1. Base case: $\epsilon \in \Sigma^*$;
   Inductive step: if $a \in \Sigma$ and $x \in \Sigma^*$ then $ax \in \Sigma^*$.
   (We will usually work with this definition.)
   Note: Recall the definition of lists!

2. Base case: $\epsilon \in \Sigma^*$;
   Inductive step: if $a \in \Sigma$ and $x \in \Sigma^*$ then $xa \in \Sigma^*$.

Note: We can prove that both definitions define exactly the same set of words.

Notation: We will often write just a instead of $a\epsilon$.

We can (recursively) *define* functions over Σ^* and
(inductively) *prove* properties about those functions.

Concatenation

Definition: Given the words x and y, the *concatenation* xy is defined as:

$$\begin{aligned} \epsilon y &= y \\ (ax')y &= a(x'y) \end{aligned}$$

Note: Recall $++$ on lists.

Example: Observe that in general $xy \neq yx$.

If $x = 010$ and $y = 11$ then $xy = 01011$ and $yx = 11010$.

Lemma: *If Σ has more than one symbol then concatenation is not commutative.*

Prefix and Suffix

Definition: Given x and y words over a certain alphabet Σ:

- x is a *prefix* of y iff there exists z such that $y = xz$;
- x is a *suffix* of y iff there exists z such that $y = zx$.

Note: $\forall\, x \in \Sigma^*.\, \epsilon$ is both a prefix and a suffix of x.

Note: $\forall\, x \in \Sigma^*.\, x$ is both a prefix and a suffix of x.

Length and Reverse

Definition: The *length* function $|\_| : \Sigma^* \to \mathbb{N}$ is defined as:

$$\begin{aligned} |\epsilon| &= 0 \\ |ax| &= 1 + |x| \end{aligned}$$

Example: $|01010| = 5$.

Definition: The *reverse* function $\mathrm{rev}(\_) : \Sigma^* \to \Sigma^*$ as:

$$\begin{aligned} \mathrm{rev}(\epsilon) &= \epsilon \\ \mathrm{rev}(ax) &= \mathrm{rev}(x)a \end{aligned}$$

Intuitively, $\mathrm{rev}(a_1 \ldots a_n) = a_n \ldots a_1$.

Note: Recall len and rev on lists.

Power

Of a string: We define x^n as follows:

$$x^0 = \epsilon$$
$$x^{n+1} = xx^n$$

Example: $(010)^3 = 010010010$.

Of an alphabet: We define Σ^n, the set of words over Σ with length n, as follows:

$$\Sigma^0 = \{\epsilon\}$$
$$\Sigma^{n+1} = \{ax \mid a \in \Sigma,\ x \in \Sigma^n\}$$

Example: $\{0,1\}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$.

Notation: $\Sigma^* = \Sigma^0 \bigcup \Sigma^1 \bigcup \Sigma^2 \ldots$ and
$\Sigma^+ = \Sigma^1 \bigcup \Sigma^2 \bigcup \Sigma^3 \ldots$

Some Properties

The following properties can be proved by induction:
(See some of the proofs on lists in slides 6–7.)

Lemma: *Concatenation is associative:* $\forall\, x, y, z \in \Sigma^*.\, x(yz) = (xy)z$.
We shall simply write xyz.

Lemma: $\forall\, x, y \in \Sigma^*.\, |xy| = |x| + |y|$.

Lemma: $\forall\, x \in \Sigma^*.\, x\epsilon = \epsilon x = x$.

Lemma: $\forall\, n \in \mathbb{N}.\, \forall\, x \in \Sigma^*.\, |x^n| = n * |x|$.

Lemma: $\forall\, n \in \mathbb{N}.\, \forall\, \Sigma.\, |\Sigma^n| = |\Sigma|^n$.

Lemma: $\forall\, x, y \in \Sigma^*.\, \text{rev}(xy) = \text{rev}(y)\text{rev}(x)$.

Lemma: $\forall\, x \in \Sigma^*.\, \text{rev}(\text{rev}(x)) = x$.

Languages

Definition: Given an alphabet Σ, a *language* $\mathcal{L}$ is a subset of Σ^*, that is, $\mathcal{L} \subseteq \Sigma^*$.

Note: If $\mathcal{L} \subseteq \Sigma^*$ and $\Sigma \subseteq \Delta$ then $\mathcal{L} \subseteq \Delta^*$.

Note: A language can be either finite or infinite.

Example: Some languages:

- Swedish, English, Spanish, French, . . . ;
- Any programming language;
- $\emptyset$, $\{\epsilon\}$ and Σ^* are languages over any Σ;
- The set of prime Natural numbers $\{1, 3, 5, 7, 11, \ldots\}$.

Some Operations on Languages

Definition: Given $\mathcal{L}$, $\mathcal{L}_1$ and $\mathcal{L}_2$ languages, we define the following languages:

Union, Intersection, ... : As for any set.

Concatenation: $\mathcal{L}_1\mathcal{L}_2 = \{x_1x_2 \mid x_1 \in \mathcal{L}_1,\ x_2 \in \mathcal{L}_2\}$.

Closure: $\mathcal{L}^* = \bigcup_{n\in\mathbb{N}} \mathcal{L}^n$ where $\mathcal{L}^0 = \{\epsilon\}$, $\mathcal{L}^{n+1} = \mathcal{L}^n\mathcal{L}$.

Note: $\emptyset^* = \{\epsilon\}$ and
$\mathcal{L}^* = \mathcal{L}^0 \cup \mathcal{L}^1 \cup \mathcal{L}^2 \cup \ldots = \{\epsilon\} \cup \{x_1 \ldots x_n \mid n > 0, x_i \in \mathcal{L}\}$

Notation: $\mathcal{L}^+ = \mathcal{L}^1 \cup \mathcal{L}^2 \cup \mathcal{L}^3 \cup \ldots$

Example: Let $\mathcal{L} = \{aa, b\}$, then
$\mathcal{L}^0 = \{\epsilon\}$, $\mathcal{L}^1 = \mathcal{L}$, $\mathcal{L}^2 = \mathcal{L}\mathcal{L} = \{aaaa, aab, baa, bb\}$, $\mathcal{L}^3 = \mathcal{L}^2\mathcal{L}$, . . .
$\mathcal{L}^* = \{\epsilon, aa, b, aaaa, aab, baa, bb, \ldots\}$.

How to Prove the Equality of Languages?

Given the languages $\mathcal{M}$ and $\mathcal{N}$, how can we prove that $\mathcal{M} = \mathcal{N}$?

A few possibilities:

- Languages are sets so we prove that $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{N} \subseteq \mathcal{M}$;
- Transitivity of equality: $\mathcal{M} = \mathcal{L}_1 = \ldots = \mathcal{L}_m = \mathcal{N}$;
- We can reason about the elements in the language:

  Example: $\{a(ba)^n \mid n \geqslant 0\} = \{(ab)^n a \mid n \geqslant 0\}$ can be proved by induction on n.

Algebraic Laws for Languages

All laws presented in slide 14 lecture 2 are valid.

In addition, we have all these laws on concatenation:

Associativity: $\mathcal{L}(\mathcal{M}\mathcal{N}) = (\mathcal{L}\mathcal{M})\mathcal{N}$

Concatenation is *not commutative:* $\mathcal{L}\mathcal{M} \neq \mathcal{M}\mathcal{L}$

Distributivity: $\mathcal{L}(\mathcal{M} \cup \mathcal{N}) = \mathcal{L}\mathcal{M} \cup \mathcal{L}\mathcal{N}$ $\quad (\mathcal{M} \cup \mathcal{N})\mathcal{L} = \mathcal{M}\mathcal{L} \cup \mathcal{N}\mathcal{L}$

Identity: $\mathcal{L}\{\epsilon\} = \{\epsilon\}\mathcal{L} = \mathcal{L}$

Annihilator: $\mathcal{L}\emptyset = \emptyset\mathcal{L} = \emptyset$

Other Rules: $\emptyset^* = \{\epsilon\}^* = \{\epsilon\}$

$\mathcal{L}^+ = \mathcal{L}\mathcal{L}^* = \mathcal{L}^*\mathcal{L}$

$(\mathcal{L}^*)^* = \mathcal{L}^*$

Algebraic Laws for Languages (Cont.)

Note: While

$$\mathcal{L}(\mathcal{M} \cap \mathcal{N}) \subseteq \mathcal{L}\mathcal{M} \cap \mathcal{L}\mathcal{N} \quad \text{and} \quad (\mathcal{M} \cap \mathcal{N})\mathcal{L} \subseteq \mathcal{M}\mathcal{L} \cap \mathcal{N}\mathcal{L}$$

both hold, in general

$$\mathcal{L}\mathcal{M} \cap \mathcal{L}\mathcal{N} \subseteq \mathcal{L}(\mathcal{M} \cap \mathcal{N}) \quad \text{and} \quad \mathcal{M}\mathcal{L} \cap \mathcal{N}\mathcal{L} \subseteq (\mathcal{M} \cap \mathcal{N})\mathcal{L}$$

don't.

Example: Consider the case where

$$\mathcal{L} = \{\epsilon, a\}, \quad \mathcal{M} = \{a\}, \quad \mathcal{N} = \{aa\}$$

Then $\mathcal{L}\mathcal{M} \cap \mathcal{L}\mathcal{N} = \{aa\}$ but $\mathcal{L}(\mathcal{M} \cap \mathcal{N}) = \mathcal{L}\emptyset = \emptyset$.

Functions between Languages

Definition: A *function* $f : \Sigma^* \to \Delta^*$ *between 2 languages* should satisfy

$$\begin{array}{l} f(\epsilon) = \epsilon \\ f(xy) = f(x)f(y) \end{array}$$

Intuitively, $f(a_1 \ldots a_n) = f(a_1) \ldots f(a_n)$.

Note: $f(a) \in \Delta^*$ if $a \in \Sigma$.

Overview of Next Lecture

Sections 2–2.2:

- DFA: deterministic finite automata.

Overview of next Week

Mon 26	Tue 27	Wed 28	Thu 29	Fri 30
		8-10 ES51 Individual help		
	Ex 10-12 EA Proofs, induction.			
Lec 13-15 HB3 DFA.				
Ex 15-17 EA Proofs, induction.		**15-17 EL41 Consultation**		

Assignment 1: Formal proofs.
Deadline: Thursday 29th March 23:59.