Automata Theory and Formal Languages
Class 1

Marcin Luckner, PhD
mluckner@mini.pw.edu.pl

Version 1.2
October 9, 2020

Major regulations - points

1. The course “Automata Theory and Formal Languages” is composed of lectures (30h) and tutorials (30h). Attendance is not obligatory.
2. During the tutorials, there will be two written tests that will check student’s knowledge of theory and her/his ability to solve practical exercises. Each test allows scoring between 0 and 50 points.
    - The tests dates will be given in future.
3. During the tutorials, there will be a chance to get extra points for solving assignments. It is possible to obtain up to 10 extra points from Formal Languages and up to 10 extra points from Automata Theory.
4. The assignments will be presented a week before classes. At the classes, a student can present his/her solution in public.

Major regulations - exams

5. After the exercises are finished, there will be an obligatory exam composed of two parts: theoretical and practical.
6. Approximately 40% of students that will achieve the best total score during a semester will be exempted one time from the practical part of the exam.
7. The exemption is not compulsory: a student, who is not satisfied with her/his score, does not have to take the exemption.

Relation

An n-ary relation $R \subseteq A_1 \times \cdots \times A_n$ is a set of n-tuples where the jth component of each n-tuple is taken from the jth domain A_j of the relation.

Binary relation

A binary relation is the special case $n = 2$ of an n-ary relation.
A binary relation over two sets A and B is a set of ordered pairs (a, b) consisting of elements a of A and elements b of B.
In a special case, $A = B$.
If R is a relation and (a, b) is a pair in R, then we note aRb.

Types of relations

Reflexive relation $(\forall a \in S)\, aRa$

Symmetric relation $(\forall a, b \in S)\, aRb \Rightarrow bRa$

Transitive relation $(\forall a, b, c \in S)\, aRb \wedge bRc \Rightarrow aRc$

Closure of relation

Suppose P is a set of properties of relation. Then P-closure of a relation R is the smallest relation R' that includes all the pairs of R and possesses the properties in P.

Equivalence relation

An *equivalence relation* is reflexive, symmetric and transitive.

Equivalence relation example I

- A word u is in a relation R with v if and only if $|u| - |v| = 3$.
  - Is that an equivalence relation?
  - If not find an equivalence closure of the relation.

Equivalence relation example II

- R is not reflexive, because $|u| - |u| = 0$.
- R is not symmetric, because $|u| - |v| = 3 \wedge |v| - |u| = -3$.
- R is not transitive, because
  $|u| - |v| = 3 \wedge |v| - |z| = 3 \wedge |u| - |z| = 6$.

Equivalence relation example III

- Relation R' $|u| - |v| = 3k, k \in \mathbb{Z}$ is an equivalence-closure of relation R.
  - R' is reflexive, for $k = 0$
  - R' is symmetric, because $|u| - |v| = 3k \wedge |v| - |u| = -3k$.
  - R' is transitive, because $|u| - |v| = 3m \wedge |v| - |z| = 3n \wedge |u| - |z| = 3(m + n)$.

Equivalence relation example IV

- A new relation is the smallest one (indirect proof)
  - Let's suppose that smaller relation R'' exists, then a pair (a, b), such that $aR'b \wedge \neg aR''b$ also exists.
  - Since $aR'b$ then $|a| - |b| = 3n$ so, exists a_1, such that $|a| - |a_1| = 3 \wedge |a_1| - |b| = 3(n-1)$.
  - Also exists a_2, such that $|a_1| - |a_2| = 3 \wedge |a_2| - |b| = 3(n-2)$.
  - Finally exists a_{n-1}, such that $|a_{n-2}| - |a_{n-1}| = 3 \wedge |a_{n-1}| - |b| = 3$.
  - That means, that a sequence of relations $aRa_1, a_1Ra_2, \ldots, a_{n-2}Ra_{n-1}, a_{n-1}Rb$ occurs.
  - If the relation R'' is a transitivity closure, then $aR''b$, and that contradicts the assumption.

Equivalence classes

Set A is an equivalence class of equivalence relation R defined on a set S if and only if

$$(\forall x, y \in A) xRy \wedge (\forall x \in A)(\forall y \notin A) \neg xRy$$

Equivalence classes example

- Name and count equivalence classes for relation R'
  $|u| - |v| = 3k, k \in \mathbb{Z}$
  - $A_0 = \{u : |u| = 3k \wedge k \in \mathbb{N}\}$
  - $A_1 = \{u : |u| = 3k + 1 \wedge k \in \mathbb{N}\}$
  - $A_2 = \{u : |u| = 3k + 2 \wedge k \in \mathbb{N}\}$
- The relation has three equivalence classes.
- However, each of the classes has the infinite number of elements.

Induction

Suppose we are given a statement $S(n)$, about an integer n to prove. One common approach is to prove two things:

1. The basis, where we show $S(i)$ for a particular integer. Usually, $i = 0$ or $i = 1$, but there are examples where we want to start at some higher level i, perhaps because the statement S is false for a few small integers.
2. The inductive step, where we assume $n \geq i$, where i is the basis integer, and we show that if $S(n)$ then $S(n + 1)$

Generalised Induction

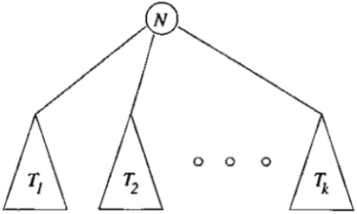
There is a more general definition

1. The basis, where we show $S(i), S(i+1), \ldots, S(j)$
2. The inductive step, if $S(i), S(i+1), \ldots, S(n)$ then $S(n+1)$

Tree

A recursive definition of the tree

1. A single node is a tree and that node is the root of the tree.
2. If $T_1, T_2, \ldots, T_k$ are trees, then we can form a new tree as follows
    2.1 Begin with a new node N, which is the root of the tree.
    2.2 Add copies of all trees $T_1, T_2, \ldots, T_k$.
    2.3 Add edges from node N to the roots of each of trees $T_1, T_2, \ldots, T_k$.



Number of leaves

Thesis: A regular tree of degree d and height n has $S(n) = d^{(n-1)}$ leaves

Proof:

1. basis: $n = 0$, a single node, one leaf $S(n) = 1$
2. inductive step:
   2.1 $S(n-1) = d^{(n-2)}$, we have $d^{(n-2)}$ leaves.
   2.2 To each old leaf we add d leaves.
   2.3 Therefore $S(n) = d^{(n-2)} * d = d^{(n-1)}$.

Alphabet and Language

- An alphabet Σ is a finite set of letters.
  - A binary alphabet $\Sigma = \{0, 1\}$
- A word is a finite sequence of letters from the alphabet.
  - $0, 101, 11111$
  - Word ϵ is a special word without letters.
- Set Σ^* contains all words that can be created over alphabet Σ.
  - $\{\epsilon, 0, 1, 00, 01, 10, \dots\}$
- Language L is a subset of Σ^*
  - $\{1, 10, 11, 100, \dots\}$
  - $\{00, 01, 10, 11\}$

Right invariant relation

A relation $R \subset \Sigma^* \times \Sigma^*$ is called a right invariant relation if and only if

$$(\forall u, v \in \Sigma^*) uRv \Rightarrow (\forall z \in \Sigma^*) uzRvz$$

Relation induced by language

A relation induced by a language L over alphabet Σ is a binary relation R_L such that:

$$(\forall u, v \in \Sigma^*) u R_L v \equiv (\forall z \in \Sigma^*) uz \in L \Leftrightarrow vz \in L$$

R_L as right invariant relation

A relation induced by a language is a right invariant relationcali.

$$(\forall u, v \in \Sigma^*)(\forall z \in \Sigma^*) uR_Lv \Rightarrow uz \in L \Leftrightarrow vz \in L$$

$$u' = uz, v = vz(\forall z' \in \Sigma^*) uR_Lv \Rightarrow u'z' \in L \Leftrightarrow v'z' \in L$$

$$(\forall u, v \in \Sigma^*) uR_Lv \Rightarrow (\forall z \in \Sigma^*) uzR_Lvz$$

R_L as equivalence relation

A relation induced by a language is an equivalence relation.

- The relation R_L is reflexive $uz \in L \Leftrightarrow uz \in L$
- The relation R_L is symmetric, $uz \in L \Leftrightarrow vz \in L$ implies $vz \in L \Leftrightarrow uz \in L$
- The relation R_L is transitive $uz \in L \Leftrightarrow vz \in L \wedge vz \in L \Leftrightarrow wz \in L$ implies $uz \in L \Leftrightarrow wz \in L$

Relation induced by language example I

A language L over alphabet $\Sigma = \{a, b\}$ is given as

$$L = \{a^m b^n : 100 \geq m \geq n \geq 1\}.$$

Find equivalence classes for the relation R_L.

Relation induced by language example II

Equivalence classes:

A_{a^i}: $\{a^i\} \wedge i = 0 \ldots 100$

A_k: $\{a^j b^i \wedge k = j - i \wedge 100 \geq j \wedge i \geq 0\} \wedge k = 0 \ldots 99$

A_R: $\{u \in \Sigma^* : (\forall z \in \Sigma^*) uz \notin L\}$

The union of classes covers Σ^*.

Relation induced by language example III

All words from the same class are in the relation R_L:

A_{a^i}: is a singleton and R_L is reflexive.

A_k: $\forall u \in A_k uz \in L \Leftrightarrow z = b^l \wedge l \leq k$

A_R: $(\forall u \in A_R)(\forall z \in \Sigma^*) uz \notin L$

Relation induced by language example IV

Words from different classes cannot be in the relation R_L:

A_{a^i}, A_{a^j}: $(\forall i < j) z = b^j \Rightarrow a^i b^j \notin L \wedge a^i b^i \in L$

A_{a^i}, A_k: if $i < 100$ then $z = ab \Rightarrow a^{i+1} b \in L \wedge a^{i+k} b^i ab \notin L$
if $i = 100$ then $z = b^{100}$

A_{a^i}, A_R: if $i = 0$ then $z = ab$
if $i > 0$ then $z = b$

A_k, A_l: $(\forall k < l) z = b^l \Rightarrow a^{i+k} b^{i+l} \notin L \wedge a^{i+l} b^{i+l} \in L$

A_k, A_R: $z = \epsilon$

Relation induced by language example V

Equivalence classes:

A_{a^i}: $\{a^i\} \wedge i = 0 \dots 100$

A_k: $\{a^j b^i \wedge k = j - i \wedge 100 \geq j \wedge i \geq 0\} \wedge k = 0 \dots 99$

A_R: $\{u \in \Sigma^* : (\forall z \in \Sigma^*) uz \notin L\}$

How many classes have we got?There are $101+100+1 = 202$ classes.

Assignments I

1. Show if the following are equivalence relations. Two words are in relationships when they are :
   1.1 Words of the same length,
   1.2 Words that consist of the same letters,
   1.3 Words that consist of the same letters in the same number
   1.4 Binary words where a difference between the number of 0s and 1s is the same
   1.5 Binary words where a remainder of division by 3 of difference between the number of 0s and 1s is the same
   1.6 Binary words where the number of sequences 111 is the same
2. Name and count equivalence classes for the equivalence relations from Assignment 1.

Assignments II

3. Name and count equivalence classes for a relation induced by language for the following languages:
   3.1 Binary words where a remainder of division by 3 of difference between the number of 0s and 1s is the same,
   3.2 Binary words with exactly one 111 sequence,
   3.3 Binary words that end with 111 sequence,
   3.4 Words that consist of all letters of the alphabet.