

# **Theory of Computation 6**

## **Homomorphisms**

**Frank Stephan**

**Department of Computer Science**

**Department of Mathematics**

**National University of Singapore**

**[fstephan@comp.nus.edu.sg](mailto:fstephan@comp.nus.edu.sg)**

# Repetition 1

Let  $(Q_1, \Sigma, \delta_1, s_1, F_1)$  and  $(Q_2, \Sigma, \delta_2, s_2, F_2)$  be dfas which recognise  $L_1$  and  $L_2$ , respectively.

Consider  $(Q_1 \times Q_2, \Sigma, \delta_1 \times \delta_2, (s_1, s_2), F)$  with  $(\delta_1 \times \delta_2)((q_1, q_2), a) = (\delta_1(q_1, a), \delta_2(q_2, a))$ . This automaton is called a **product automaton** and one can choose  $F$  such that it recognises the union or intersection or difference of the respective languages.

Union:  $F = F_1 \times Q_2 \cup Q_1 \times F_2$ ;

Intersection:  $F = F_1 \times F_2 = F_1 \times Q_2 \cap Q_1 \times F_2$ ;

Difference:  $F = F_1 \times (Q_2 - F_2)$ ;

Symmetric Difference:  $F = F_1 \times (Q_2 - F_2) \cup (Q_1 - F_1) \times F_2$ .

# Repetition 2 and Gaps Filled

Regular languages are also closed under Kleene star, Kleene plus and concatenation: Use nfas for these and convert to dfas.

Context-free languages are closed under union, Kleene star, Kleene plus, concatenation and intersection with regular languages. They are in general not closed under intersection and complement.

Context-sensitive languages are closed under union, intersection, Kleene star, Kleene plus and concatenation. While these are easy to see, the following result is more difficult: They are also closed under complement (not part of this course).

Recursively enumerable languages are closed under union, intersection, Kleene star, Kleene plus and concatenation; they are not closed under complement.

# Repetition 3: Palindromes

The members of the language  $\{x \in \Sigma^* : x = x^{mi}\}$  are called palindromes. A palindrome is a word or phrase which looks the same from both directions.

An example is the German name “OTTO”; furthermore, when ignoring spaces and punctuation marks, a famous palindrome is the phrase “A man, a plan, a canal: Panama.” originating from the time when the canal in Panama was built.

The grammar with the rules  $S \rightarrow aSa|aa|a|\epsilon$  with  $a$  ranging over all members of  $\Sigma$  generates all palindromes; so for  $\Sigma = \{0, 1, 2\}$  the rules of the grammar would be  $S \rightarrow 0S0 | 1S1 | 2S2 | 00 | 11 | 22 | 0 | 1 | 2 | \epsilon$ .

The set of palindromes is not regular.

# Homomorphism

## Example

Let  $\text{ascii}(\text{Year 2019}) = 596561722032303139$  represent each letter of “Year 2019” by its two-digit hexadecimal ASCII representation.

## Definition 6.1

A homomorphism is a mapping  $h$  with domain  $\Sigma^*$  for some alphabet  $\Sigma$  which preserves concatenation:

$$h(\mathbf{v} \cdot \mathbf{w}) = h(\mathbf{v}) \cdot h(\mathbf{w}).$$

## Proposition 6.2

The homomorphism is determined by the images of the single letters and  $h(\mathbf{w}) = h(\mathbf{a}_1) \cdot h(\mathbf{a}_2) \cdot \dots \cdot h(\mathbf{a}_n)$  for a word  $\mathbf{w} = \mathbf{a}_1\mathbf{a}_2 \dots \mathbf{a}_n$ ;  $h(\varepsilon) = \varepsilon$ .

## Quiz

What is  $\text{ascii}(\text{Year 1819})$  for above homomorphism  $\text{ascii}$ ?

# Exercises 6.3 and 6.4

Count the number of homomorphisms and list them; explain why there are not more. Two homomorphisms are the same iff they have the same values  $h(0)$ ,  $h(1)$ ,  $h(2)$ ,  $h(3)$ . Here they take values from  $4^*$ .

## Exercise 6.3

How many homomorphisms  $h$  satisfy  $h(012) = 44444$ ,  $h(102) = 444444$ ,  $h(00) = 44444$  and  $h(3) = 4$ ?

## Exercise 6.4

How many homomorphisms  $h$  satisfy  $h(012) = 44444$ ,  $h(102) = 44444$ ,  $h(0011) = 444444$  and  $h(3) = 44$ ?

# Homomorphic Images

## Theorem 6.5

The homomorphic images of regular and context-free languages are regular and context-free, respectively.

## Construction

Given a homomorphism  $h$ , replace in any rule of a given regular / context-free grammar every terminal  $a$  by the word  $h(a)$ ; these replacements only occur on the right side of the rules. The type of the grammar remains unchanged.

For a proof that  $S \Rightarrow^* w$  in the original grammar iff  $S \Rightarrow h(w)$  in the new grammar, one shows by induction for a derivation  $S \Rightarrow v_1 \Rightarrow \dots \Rightarrow v_n \Rightarrow w$  translates into  $h(S) \Rightarrow h(v_1) \Rightarrow \dots \Rightarrow h(v_n) \Rightarrow h(w)$  where  $h$  is extended by letting  $h(A) = A$  for all non-terminals  $A$ . The converse also holds.

# Example 6.6

One can apply the homomorphisms also directly to regular expressions using the rules  $\mathbf{h}(\mathbf{L} \cup \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cup \mathbf{h}(\mathbf{H})$ ,  $\mathbf{h}(\mathbf{L} \cdot \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cdot \mathbf{h}(\mathbf{H})$  and  $\mathbf{h}(\mathbf{L}^*) = (\mathbf{h}(\mathbf{L}))^*$ . Thus one can move a homomorphism into the inner parts (which are the finite sets used in the regular expression) and then apply the homomorphism there.

So for the language  $(\{0, 1\}^* \cup \{0, 2\}^*) \cdot \{33\}^*$  and the homomorphism which maps each symbol  $\mathbf{a}$  to  $\mathbf{aa}$ , one obtains the language  $(\{00, 11\}^* \cup \{00, 22\}^*) \cdot \{3333\}^*$ .



# Homomorphisms and Growth

## Exercise 6.7

Consider the following statements for regular languages  $L$ :

- (a)  $h(\emptyset) = \emptyset$ ;
- (b) If  $L$  is finite so is  $h(L)$ ;
- (c) If  $L$  has polynomial growth so has  $h(L)$ ;
- (d) If  $L$  has exponential growth so has  $h(L)$ .

Which of these statements are true and which are false? Prove the answers. Use the following rules: Example 6.6;  $H^*$  has polynomial growth iff  $H \subseteq \{u\}^*$  for some word  $u$ ; if  $H, K$  have polynomial growth so do  $H \cup K$  and  $H \cdot K$ .

## Exercise 6.8

Construct a context-sensitive language  $L$  and a homomorphism  $h$  such that  $L$  has polynomial growth and  $h(L)$  has exponential growth.

# Homomorphism Reduce Kleene star

One can reduce the number of stars in  $\bigcup_{a \in \Sigma} aa^*$  to two using intersection:

$$\begin{aligned} &00^* \cup 11^* \cup 22^* \cup 33^* = \\ &(\{0, 1, 2, 3\} \cdot \{00, 11, 22, 33\}^* \cdot \{\varepsilon, 0, 1, 2, 3\}) \cap \\ &(\{00, 11, 22, 33\}^* \cdot \{\varepsilon, 0, 1, 2, 3\}). \end{aligned}$$

The general result needs also a homomorphism.

## Theorem 6.9

Let  $L$  be a regular language. Then there are two regular expressions  $\sigma, \tau$  each containing only one Kleene star and some finite sets and concatenations and there is one homomorphism  $h$  such that  $L$  is described by  $h(\sigma \cap \tau)$ .

The idea is to encode states of a dfa into the symbols; expressions  $\sigma$  and  $\tau$  test state-transitions at even and odd positions, respectively;  $h$  removes the state markers from the symbols.

# Construction

Let  $(Q, \Sigma, \delta, s, F)$  be a dfa recognising the language and let  $\Gamma = Q \times \Sigma$  and

$$\Gamma_1 = \{(q, a)(p, b) \in \Gamma \times \Gamma : \delta(q, a) = p\};$$

$$\Gamma_2 = \{(q, a)(p, b) \in \Gamma_1 : \delta(p, b) \in F\};$$

$$\Gamma_3 = \{(q, a) : \delta(q, a) \in F\};$$

$$\Gamma_4 = \{\varepsilon : s \in F\};$$

$$\Gamma_5 = \{(s, a) : a \in \Sigma\}.$$

The expression is  $\mathbf{h}(\sigma \cap \tau)$  where  $\mathbf{h}((q, a)) = a$ ;

$$\sigma = (\Gamma_1^* \cdot (\Gamma_2 \cup \Gamma_3) \cup \Gamma_4);$$

$$\tau = (\Gamma_5 \cdot \Gamma_1^* \cdot (\Gamma \cup \{\varepsilon\}) \cup \Gamma_4).$$

Odd transitions and acceptance checked by  $\sigma$ ;

Even transitions and start checked by  $\tau$ .

# Context-Sensitive Languages

## Theorem 6.11

Every recursively enumerable language (= language generated by some grammar) is the homomorphic image of a context-sensitive language.

The idea is that if some grammar generates  $(\mathbf{N}, \{1, 2, \dots, k\}, \mathbf{P}, \mathbf{S})$  for  $\mathbf{L}$ , one can make a new grammar for a context-sensitive language  $\mathbf{H}$  such that for all  $w \in \{1, 2, \dots, k\}^*$ ,  $w \in \mathbf{L}$  iff  $w \cdot 0^\ell \in \mathbf{H}$  for some  $\ell$ . These additional  $0$  will be used to make words longer so that in the new grammar, all rules  $\mathbf{l} \rightarrow \mathbf{r}$  satisfy  $|\mathbf{l}| \leq |\mathbf{r}|$  which is obtained sufficiently many  $0$  on the right side and by making rules for  $0$  to swap with other symbols to move right.

# Images of Homomorphisms

Determine  $h(L)$  for the following languages:

(a)  $\{0, 1, 2\}^*$ ;

(b)  $\{00, 11, 22\}^* \cap \{000, 111, 222\}^*$ ;

(c)  $(\{00, 11\}^* \cup \{00, 22\}^* \cup \{11, 22\}^*) \cdot \{011222\}$ ;

(d)  $\{w \in \{0, 1\}^* : w \text{ has more 1s than it has 0s}\}$ .

Exercise 6.13

$h$  is given as  $h(0) = 1$ ,  $h(1) = 22$ ,  $h(2) = 333$ .

Exercise 6.14

$h$  is given as  $h(0) = 3$ ,  $h(1) = 4$ ,  $h(2) = 334433$ .

# Exercise 6.15

Let a homomorphism  $h : \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}^* \rightarrow \{0, 1, 2, 3\}^*$  be given by the equations  $h(0) = 0$ ,  $h(1) = h(4) = h(7) = 1$ ,  $h(2) = h(5) = h(8) = 2$ ,  $h(3) = h(6) = h(9) = 3$ . Interpret the images of  $h$  as quarternary numbers (numbers of base four, so **12321** represents **1** times two hundred fifty six plus **2** times sixty four plus **3** times sixteen plus **2** times four plus **1**). Prove the following:

- Every quarternary number is the image of a decimal number without leading zeroes;
- A decimal number  $w$  has leading zeroes iff the quarternary number  $h(w)$  has leading zeroes;
- A decimal number  $w$  is a multiple of three iff the quarternary number is a multiple of three.

# Exercise 6.16

Consider only homomorphisms

$h : \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}^* \rightarrow \{0, 1\}^*$  such that

- $h(w)$  has leading zeroes iff  $w$  has;
- $h(0) = 0$ ;
- the range of  $h$  is  $\{0, 1\}^*$ .

For each of  $p = 2, 3, 5$ , answer the following question: Can one choose  $h$  such that, in addition,  $w$  is a multiple of  $p$  iff  $h(w)$  is as a binary number, is a multiple of  $p$ ?

If  $h$  can be chosen as desired then list this  $h$  else prove that such a homomorphism  $h$  cannot exist.

# Exercise 6.17

Construct a homomorphism

$h : \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}^* \rightarrow \{0, 1\}^*$  such that for every  $w$  the number  $h(w)$  has never leading zeroes and the remainder of the decimal number  $w$  when divided by nine is the same as the remainder of the binary number  $h(w)$  when divided by nine.

Note that here it is not required that the range covers all binary numbers.



# Fibonacci Representation

Let  $a_0 = 1$ ,  $a_1 = 1$ ,  $a_2 = 2$  and, for all  $n$ ,  $a_{n+2} = a_n + a_{n+1}$ . Every number is the sum of non-neighbouring Fibonacci numbers: For each non-zero  $n$  there is a unique  $b_m b_{m-1} \dots b_0 \in (10^+)^+$  with

$$n = \sum_{k=0,1,\dots,m} b_k \cdot a_k.$$

So **1010** represents four and **100100** represents ten.

**Exercise 6.18:** Construct a homomorphism

$h : \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \rightarrow \{0, 1\}^*$  such that  $h(0) = 0$  and the image of all decimal numbers is the regular set

$\{0\} \cup (10^+)^+$ . Show that all  $h$  satisfying this also satisfy the

following statement: For every  $p > 1$  there is a decimal number  $w$  such that ( $w$  is a multiple of  $p$  iff  $h(w)$  is not a multiple of  $p$ ).

# Inverse Homomorphism

## Description 6.19

Let  $h$  have domain  $\Sigma^*$  and the set

$h^{-1}(L) = \{w \in \Sigma^* : h(w) \in L\}$  is called the inverse image of

$h$ .  $h^{-1}$  satisfies the following rules:

(a)  $h^{-1}(L) \cap h^{-1}(H) = h^{-1}(L \cap H)$ ;

(b)  $h^{-1}(L) \cup h^{-1}(H) = h^{-1}(L \cup H)$ ;

(c)  $h^{-1}(L) \cdot h^{-1}(H) \subseteq h^{-1}(L \cdot H)$ ;

(d)  $h^{-1}(L)^* \subseteq h^{-1}(L^*)$ .

# Theorem 6.20 and Exercise 6.21

## Theorem 6.20

If  $\mathbf{L}$  is on level  $\mathbf{k}$  of the Chomsky hierarchy and  $\mathbf{h}$  is an homomorphism then  $\mathbf{h}^{-1}(\mathbf{L})$  is on level  $\mathbf{k}$  of the Chomsky hierarchy.

Construction for the regular case: If  $(\mathbf{Q}, \mathbf{\Gamma}, \gamma, \mathbf{s}, \mathbf{F})$  is a dfa recognising  $\mathbf{L}$  and  $\mathbf{h} : \Sigma^* \rightarrow \Gamma^*$  is an homomorphism then  $(\mathbf{Q}, \Sigma, \delta, \mathbf{s}, \mathbf{F})$  is a dfa recognising  $\mathbf{h}^{-1}(\mathbf{L})$  where, for every  $\mathbf{q} \in \mathbf{Q}$  and  $\mathbf{a} \in \Sigma$ ,  $\delta(\mathbf{q}, \mathbf{a}) = \gamma(\mathbf{q}, \mathbf{h}(\mathbf{a}))$ .

## Exercise 6.21

Let  $\mathbf{h} : \{0, 1, 2, 3\}^* \rightarrow \{0, 1, 2, 3\}^*$  be given by  $\mathbf{h}(0) = 00$ ,  $\mathbf{h}(1) = 012$ ,  $\mathbf{h}(2) = 123$  and  $\mathbf{h}(3) = 1$  and let  $\mathbf{L}$  consist of all words containing exactly five 0s and at least one 2.

Construct a complete dfa recognising  $\mathbf{h}^{-1}(\mathbf{L})$ .

# Generalised Homomorphism

## Description 6.22

A **generalised homomorphism** is a mapping from regular sets to regular sets which satisfies  $\mathbf{h}(\mathbf{L} \cup \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cup \mathbf{h}(\mathbf{H})$ ,  $\mathbf{h}(\mathbf{L} \cdot \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cdot \mathbf{h}(\mathbf{H})$ ,  $\mathbf{h}(\mathbf{L}^*) = (\mathbf{h}(\mathbf{L}))^*$  and  $\mathbf{h}(\emptyset) = \emptyset$  for all regular sets  $\mathbf{L}$  and  $\mathbf{H}$ .

## Examples 6.23

The following mappings are generalised homomorphisms:

- $\mathbf{L} \mapsto \mathbf{L} \cap \{\varepsilon\}$ ;
- $\emptyset \mapsto \emptyset$  and  $\mathbf{L} \mapsto \{\varepsilon\}$  for all non-empty sets  $\mathbf{L}$ ;
- $\emptyset \mapsto \emptyset$ ,  $\{\varepsilon\} \mapsto \{\varepsilon\}$  and  $\mathbf{L} \mapsto \Sigma^*$  for all other sets  $\mathbf{L}$ ;
- $\mathbf{L} \mapsto \mathbf{L}$  (identity mapping);
- $\mathbf{L} \mapsto \{\mathbf{v} \in \Sigma^* : \exists \mathbf{w} \in \mathbf{L} [|\mathbf{v}| = |\mathbf{w}|]\}$ .

# Exercises 6.24-6.25

## Exercise 6.24

Show that whenever  $h : \Sigma^* \rightarrow \Gamma^*$  is a homomorphism then the mapping  $L \mapsto \{h(u) : u \in L\}$  is a generalised homomorphism which maps regular subsets of  $\Sigma^*$  to regular subsets of  $\Gamma^*$ .

## Exercise 6.25

Let  $h$  be any given generalised homomorphism. Show by structural induction that  $h(L) = \bigcup_{u \in L} h(u)$  for all regular languages  $L$ . Furthermore, show that every mapping  $h$  satisfying  $h(\{\varepsilon\}) = \{\varepsilon\}$ ,  $h(L) = \bigcup_{u \in L} h(\{u\})$  and  $h(L \cdot H) = h(L) \cdot h(H)$  for all regular subsets  $L, H$  of  $\Sigma^*$  is a generalised homomorphism. Is the same true if one weakens the condition  $h(\{\varepsilon\}) = \{\varepsilon\}$  to  $\varepsilon \in h(\{\varepsilon\})$ ?

# Exercises 6.26-6.28

## Exercise 6.26

Construct a mapping which satisfies  $\mathbf{h}(\emptyset) = \emptyset$ ,  $\mathbf{h}(\{\varepsilon\}) = \{\varepsilon\}$ ,  $\mathbf{h}(\mathbf{L} \cup \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cup \mathbf{h}(\mathbf{H})$  and  $\mathbf{h}(\mathbf{L} \cdot \mathbf{H}) = \mathbf{h}(\mathbf{L}) \cdot \mathbf{h}(\mathbf{H})$  for all regular languages  $\mathbf{L}, \mathbf{H}$  but which does not satisfy  $\mathbf{h}(\mathbf{L}) = \bigcup_{\mathbf{u} \in \mathbf{L}} \mathbf{h}(\{\mathbf{u}\})$  for some infinite regular set  $\mathbf{L}$ .

## Exercise 6.27

Assume that  $\mathbf{h}$  is a generalised homomorphism and  $\mathbf{k}(\mathbf{L}) = \mathbf{h}(\mathbf{L}) \cdot \mathbf{h}(\mathbf{L})$ . Is  $\mathbf{k}$  a generalised homomorphism? Prove the answer.

## Exercise 6.28

Assume that  $\mathbf{h}$  is a generalised homomorphism and  $\ell(\mathbf{L}) = \bigcup_{\mathbf{u} \in \mathbf{h}(\mathbf{L})} \Sigma^{|\mathbf{u}|}$ , where  $\Sigma^0 = \{\varepsilon\}$ . Is  $\ell$  a generalised homomorphism? Prove the answer.

# Exercise 6.29

Let  $\Sigma = \{0, 1, 2\}$  and  $h$  be the generalised homomorphism given by  $h(\{0\}) = \{1, 2\}$ ,  $h(\{1\}) = \{0, 2\}$  and  $h(\{2\}) = \{0, 1\}$ . Which of the following statements are true for this  $h$  and all regular subsets  $L, H$  of  $\Sigma^*$ :

- (a) If  $L \neq H$  then  $h(L) \neq h(H)$ ;
- (b) If  $L \subseteq H$  then  $h(L) \subseteq h(H)$ ;
- (c) If  $L$  is finite then  $h(L)$  is finite;
- (d) If  $L$  is infinite then  $h(L)$  is infinite and has exponential growth.

Prove the answers. The formula  $h(L) = \bigcup_{u \in L} h(\{u\})$  from Exercise 6.25 can be used without proof for this exercise.

# Generalised Homomorphisms

Determine  $h(L) = \bigcup_{a_1 a_2 \dots a_n \in L} h(a_1) \cdot h(a_2) \cdot \dots \cdot h(a_n)$  for the following languages; if possible give regular expressions.

(a)  $\{00, 01, 02, 10, 11, 12, 20, 21, 22\}^*$ ;

(b)  $\{00, 11, 22\}^* \cdot \{000, 111, 222\}$ ;

(c)  $\{0^n 1^n 2^n : n \geq 2\}$ .

## Exercise 6.30

$h$  is given as  $h(0) = \{3, 4\}^+$ ,  $h(1) = \{3, 5\}^+$ ,  $h(2) = \{4, 5\}^+$ .

## Exercise 6.31

$h$  is given as  $h(0) = \{\varepsilon, 3, 33\}$ ,  $h(1) = \{\varepsilon, 4, 44\}$ ,  
 $h(2) = \{\varepsilon, 5, 55\}$ .

## Exercise 6.32

$h$  is given as  $h(a) = \{aaa, aaaa\}^+$  for all letters  $a \in \{0, 1, 2\}$ .