Tomi Kärki

Compatibility relations on codes and free monoids

University of Turku and Turku Centre for Computer Science

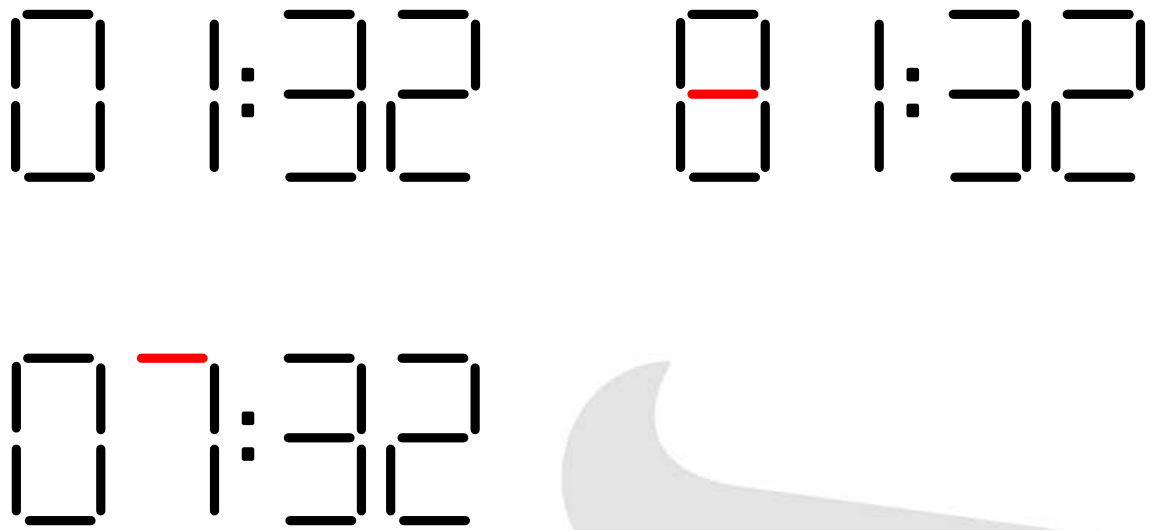
(TUCS)

Introduction

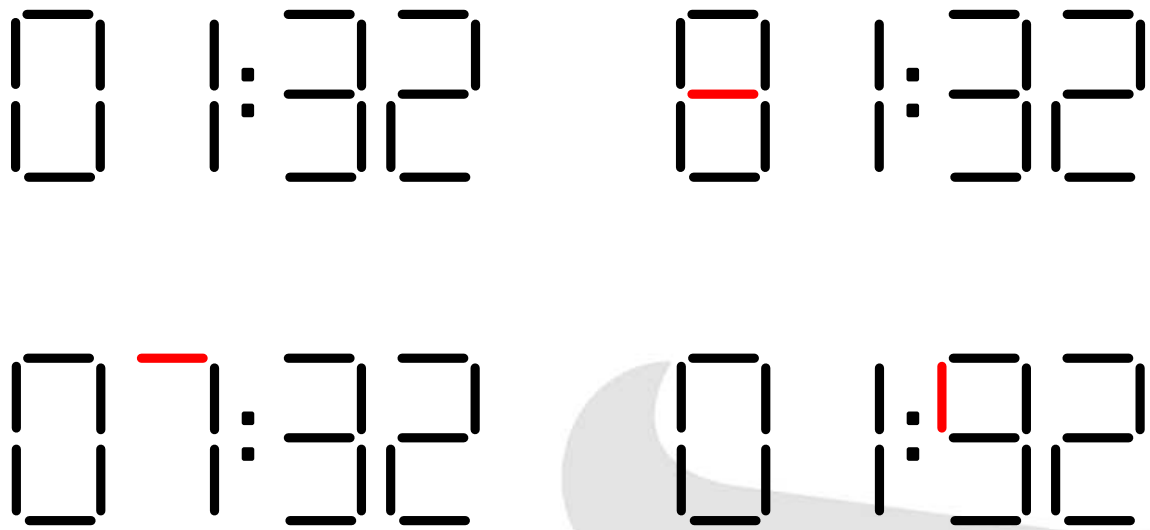
01:32

Introduction

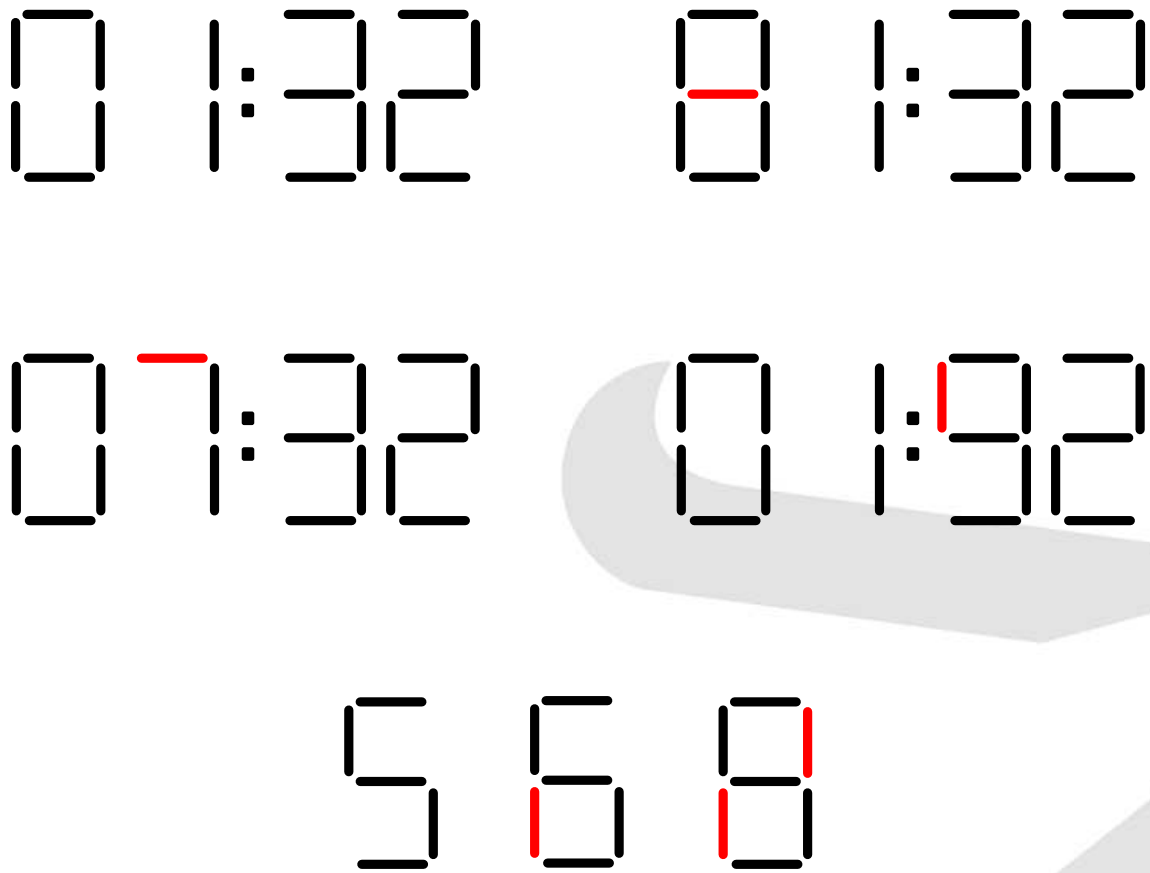
Introduction



Introduction



Introduction



Outline of Topics

- Word relations
- Relational codes
- Minimal and maximal relations
- Relationally free monoids and stability
- Hulls
- Defect effect

Notations

A	an alphabet
ε	empty word
X	a set of words over A^*
$R \subseteq X \times X$	a relation on X
$x\,R\,y$	$(x, y) \in R$
ι_X	$\{(x, x) \mid x \in X\}$
Ω_X	$\{(x, y) \mid x, y \in X\}$
$\langle R \rangle$	reflexive and symmetric closure of R
R_Y	$R \cap (Y \times Y)$
$R(X)$	$\{u \in A^* \mid \exists\, x \in X :\ x\,R\,u\}$

Word relations

- *compatibility relation* = reflexive and symmetric

Word relations

- *compatibility relation* = reflexive and symmetric
- *word relation* R = compatibility relation and

$$a_1 \cdots a_m \, R \, b_1 \cdots b_n \Leftrightarrow m = n \text{ and } a_i \, R \, b_i \text{ for all } i = 1, 2, \ldots, m$$

where $a_1, \ldots, a_m, b_1, \ldots, b_n \in A$

Word relations

- *compatibility relation* = reflexive and symmetric
- *word relation* R = compatibility relation and

  $a_1 \cdots a_m \, R \, b_1 \cdots b_n \Leftrightarrow m = n$ and $a_i \, R \, b_i$ for all $i = 1, 2, \ldots, m$

  where $a_1, \ldots, a_m, b_1, \ldots, b_n \in A$
- If $u \, R \, v$, then words u and v are *R-compatible*

Word relations

- *compatibility relation* = reflexive and symmetric
- *word relation* R = compatibility relation and

$$a_1 \cdots a_m \, R \, b_1 \cdots b_n \Leftrightarrow m = n \text{ and } a_i \, R \, b_i \text{ for all } i = 1, 2, \ldots, m$$

where $a_1, \ldots, a_m, b_1, \ldots, b_n \in A$

- If $u \, R \, v$, then words u and v are *R-compatible*
- $\begin{cases} \text{multiplicativity:} & u \, R \, v, u' \, R \, v' \Rightarrow uu' \, R \, vv', \\ \text{simplifiability:} & uu' \, R \, vv', |u| = |v| \Rightarrow u \, R \, v, \; u' \, R \, v' \end{cases}$

Word relations

Example 1. $A = \{a, b, c\}$
$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

Word relations

Example 1. $A = \{a, b, c\}$

$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

$abba \, R \, baab$

Word relations

Example 1. $A = \{a, b, c\}$

$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

$abba \, R \, baab$

$abc \not R \, cbc$

Word relations

Example 1.

$$A = \{a, b, c\}$$
$$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$$

$$abba \, R \, baab$$
$$abc \not R \, cbc$$

Example 2.

Word relations

Example 1. $A = \{a, b, c\}$

$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

$abba \, R \, baab$

$abc \not R \, cbc$

Example 2. Partial words

Word relations

Example 1. $A = \{a, b, c\}$

$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

$abba \, R \, baab$

$abc \not R \, cbc$

Example 2. Partial words

$k\,n\diamondsuit w\,l\,\diamondsuit\,dg\,e$

$\diamondsuit n\,o\,w\diamondsuit\diamondsuit dg\diamondsuit$

$k\,n\,o\,w\,l\;e\;dg\,e$

Word relations

Example 1. $A = \{a, b, c\}$

$R = \langle\{(a, b)\}\rangle = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$

$abba \, R \, baab$

$abc \not R \, cbc$

Example 2. Partial words

$k\,n\diamondsuit w\,l\,\diamondsuit\,dg\,e$

$\diamondsuit n\,o\,w\diamondsuit\diamondsuit dg\diamondsuit$

$k\,n\,o\,w\,l\;e\;dg\,e$

$$R_{\uparrow} = \langle\{(\diamondsuit, a) \mid a \in A\}\rangle$$

Relational codes

- Let R and S be word relations

Relational codes

- Let R and S be word relations
- $X \subseteq A^*$ is an (R, S)-*code* if for all $n, m \geq 1$ and $x_1, \ldots, x_m, y_1, \ldots, y_n \in X$, we have

$$x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m$$

Relational codes

- Let R and S be word relations
- $X \subseteq A^*$ is an (R, S)-*code* if for all $n, m \geq 1$ and $x_1, \ldots, x_m, y_1, \ldots, y_n \in X$, we have

$$\boxed{x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m}$$

- (R, S)-code *relational* code
  (R, ι)-code *strong* R-code
  (R, R)-code *weak* R-code
  (ι, ι)-code code

Relational codes

Example. $A = \{a, b, c\}$
$X = \{ab, c\}$
$S = \iota$

$R = \iota$	
$R = \langle\{(a, c)\}\rangle$	
$R = \langle\{(a, c), (b, c)\}\rangle$	

Relational codes

Example. $A = \{a, b, c\}$
$X = \{ab, c\}$
$S = \iota$

$R = \iota$	(prefix) code
$R = \langle\{(a, c)\}\rangle$	
$R = \langle\{(a, c), (b, c)\}\rangle$	

Relational codes

Example. $A = \{a, b, c\}$
$X = \{ab, c\}$
$S = \iota$

$R = \iota$	(prefix) code
$R = \langle\{(a, c)\}\rangle$	(R, ι)-code
$R = \langle\{(a, c), (b, c)\}\rangle$	

Relational codes

Example. $A = \{a, b, c\}$
$X = \{ab, c\}$
$S = \iota$

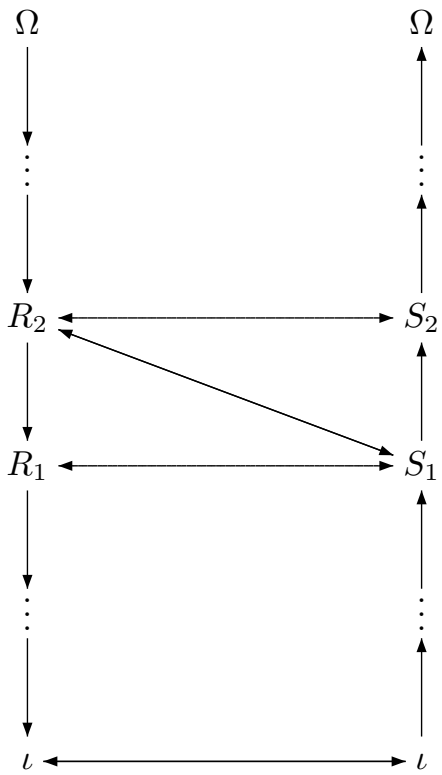
$R = \iota$	(prefix) code
$R = \langle\{(a, c)\}\rangle$	(R, ι)-code
$R = \langle\{(a, c), (b, c)\}\rangle$	$ab\, R\, c.c$

Relational codes

$$x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m$$

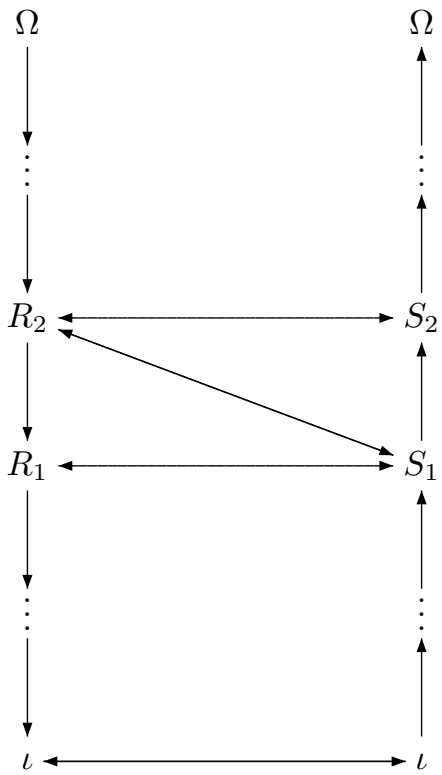
Relational codes

$$x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m$$



Relational codes

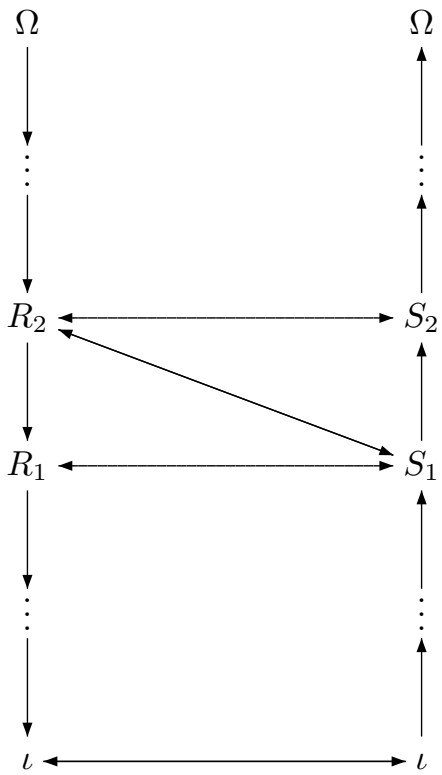
$$x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m$$



Theorem 3. Every (R, S)-code X is a code.

Relational codes

$$x_1 \cdots x_m \, R \, y_1 \cdots y_n \Rightarrow n = m \text{ and } x_i \, S \, y_i \text{ for } i = 1, 2, \ldots, m$$



Theorem 3. Every (R, S)-code X is a code.

Theorem 4. Let X be a subset of A^*. X is an (R, S)-code $\Leftrightarrow$ X is an (R, R)-code and $R_X \subseteq S_X$.

Modified Sardinas-Patterson algorithm

Modified Sardinas-Patterson algorithm

- finite $X \subseteq A^+$

Modified Sardinas-Patterson algorithm

- finite $X \subseteq A^+$
- $U_1 = R(X)^{-1}X \setminus \{\varepsilon\}$

Modified Sardinas-Patterson algorithm

- finite $X \subseteq A^+$
- $U_1 = R(X)^{-1}X \setminus \{\varepsilon\}$
- $U_{n+1} = R(X)^{-1}U_n \ \cup \ R(U_n)^{-1}X$ for $n \geq 1$

Modified Sardinas-Patterson algorithm

- finite $X \subseteq A^+$
- $U_1 = R(X)^{-1}X \setminus \{\varepsilon\}$
- $U_{n+1} = R(X)^{-1}U_n \ \cup \ R(U_n)^{-1}X$ for $n \geq 1$
- Let $i \geq 2$ satisfy $U_i = U_{i-t}$ for some $t > 0$

Modified Sardinas-Patterson algorithm

- finite $X \subseteq A^+$
- $U_1 = R(X)^{-1}X \setminus \{\varepsilon\}$
- $U_{n+1} = R(X)^{-1}U_n \ \cup \ R(U_n)^{-1}X$ for $n \geq 1$
- Let $i \geq 2$ satisfy $U_i = U_{i-t}$ for some $t > 0$
- X is a weak R-code if and only if

$$\boxed{\varepsilon \notin \bigcup_{j=1}^{i-1} U_j}$$

Modified Sardinas-Patterson algorithm

Example.

$$A = \{a, b, c\}$$
$$X = \{abb, ca, c\}$$
$$R = \langle\{(a, b), (b, c)\}\rangle$$

Modified Sardinas-Patterson algorithm

Example.

$$A = \{a, b, c\}$$
$$X = \{abb, ca, c\}$$
$$R = \langle\{(a, b), (b, c)\}\rangle$$
$$U_1 = R(X)^{-1}X \setminus \{\varepsilon\} = \{a\}$$

Modified Sardinas-Patterson algorithm

Example.

$$A = \{a, b, c\}$$
$$X = \{abb, ca, c\}$$
$$R = \langle\{(a, b), (b, c)\}\rangle$$

$$U_1 = R(X)^{-1}X \setminus \{\varepsilon\} = \{a\}$$
$$U_2 = R(X)^{-1}U_1 \ \cup \ R(U_1)^{-1}X = \emptyset \ \cup \ \{bb\}$$

Modified Sardinas-Patterson algorithm

Example.

$$A = \{a, b, c\}$$
$$X = \{abb, ca, c\}$$
$$R = \langle\{(a, b), (b, c)\}\rangle$$

$$U_1 = R(X)^{-1}X \setminus \{\varepsilon\} = \{a\}$$
$$U_2 = R(X)^{-1}U_1 \ \cup \ R(U_1)^{-1}X = \emptyset \ \cup \ \{bb\}$$
$$U_3 = R(X)^{-1}U_2 \ \cup \ R(U_2)^{-1}X = \{\varepsilon, b\} \ \cup \ \{\varepsilon, b\}$$

Modified Sardinas-Patterson algorithm

Example.

$$A = \{a, b, c\}$$
$$X = \{abb, ca, c\}$$
$$R = \langle\{(a, b), (b, c)\}\rangle$$

$$U_1 = R(X)^{-1}X \setminus \{\varepsilon\} = \{a\}$$
$$U_2 = R(X)^{-1}U_1 \ \cup \ R(U_1)^{-1}X = \emptyset \ \cup \ \{bb\}$$
$$U_3 = R(X)^{-1}U_2 \ \cup \ R(U_2)^{-1}X = \{\varepsilon, b\} \ \cup \ \{\varepsilon, b\}$$

$\Longrightarrow$ X is not an (R, R)-code

$$ca.ca \, R \, c.abb$$

Minimal and maximal relations

$S \in S_{\min}(X, R)$: X is an (R, S)-code

$\forall S' \subset S : X$ is not an (R, S')-code

Minimal and maximal relations

$S \in S_{\min}(X, R)$: X is an (R, S)-code
$\forall S' \subset S$: X is not an (R, S')-code

$S \in S_{\max}(X, R)$: X is an (R, S)-code
$\forall S' \supset S$: X is not an (R, S')-code

$R \in R_{\min}(X, S)$: X is an (R, S)-code
$\forall R' \subset R$: X is not an (R', S)-code

$R \in R_{\max}(X, S)$: X is an (R, S)-code
$\forall R' \supset R$: X is not an (R', S)-code

Minimal and maximal relations

$S \in S_{\min}(X, R)$: X is an (R, S)-code
$\forall S' \subset S : X$ is not an (R, S')-code

$S \in S_{\max}(X, R)$: X is an (R, S)-code
$\forall S' \supset S : X$ is not an (R, S')-code

$R \in R_{\min}(X, S)$: X is an (R, S)-code
$\forall R' \subset R : X$ is not an (R', S)-code

$R \in R_{\max}(X, S)$: X is an (R, S)-code
$\forall R' \supset R : X$ is not an (R', S)-code

- $S_{\max}(X, R) = \{\Omega\}$

Minimal and maximal relations

$S \in S_{\min}(X, R)$: X is an (R, S)-code
$\forall S' \subset S : X$ is not an (R, S')-code

$S \in S_{\max}(X, R)$: X is an (R, S)-code
$\forall S' \supset S : X$ is not an (R, S')-code

$R \in R_{\min}(X, S)$: X is an (R, S)-code
$\forall R' \subset R : X$ is not an (R', S)-code

$R \in R_{\max}(X, S)$: X is an (R, S)-code
$\forall R' \supset R : X$ is not an (R', S)-code

- $S_{\max}(X, R) = \{\Omega\}$
- $R_{\min}(X, S) = \{\iota\}$

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element
- finding $S_{\min}(X, R)$ easy

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element
- finding $S_{\min}(X, R)$ easy
- $R_{\max}(X, S)$ can contain relations of different size

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element
- finding $S_{\min}(X, R)$ easy
- $R_{\max}(X, S)$ can contain relations of different size
- finding $R_{\max}(X, S)$ hard for arbitrary alphabets

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element
- finding $S_{\min}(X, R)$ easy
- $R_{\max}(X, S)$ can contain relations of different size
- finding $R_{\max}(X, S)$ hard for arbitrary alphabets

Problem:	MAXIMAL RELATION
Instance:	$X \subseteq A^+$, relation S, $k \in \mathbb{N}$
Question:	Is max. size of $R \in R_{\max}(X, S) \geq k$?

Minimal and maximal relations

- $S_{\min}(X, R)$ is a unique element
- finding $S_{\min}(X, R)$ easy
- $R_{\max}(X, S)$ can contain relations of different size
- finding $R_{\max}(X, S)$ hard for arbitrary alphabets

Problem: MAXIMAL RELATION
Instance: $X \subseteq A^+$, relation S, $k \in \mathbb{N}$
Question: Is max. size of $R \in R_{\max}(X, S) \geq k$?

NP-complete

Relationally free monoids

A monoid $M \subseteq A^*$ is (R, S)-*free* if it has a subset $B \subseteq M$ (called an (R, S)-base of M) such that

$$(i) \quad M = B^*,$$

$$(ii) \quad B \text{ is an } (R, S)\text{-code.}$$

Relationally free monoids

A monoid $M \subseteq A^*$ is (R, S)-*free* if it has a subset $B \subseteq M$ (called an (R, S)-base of M) such that

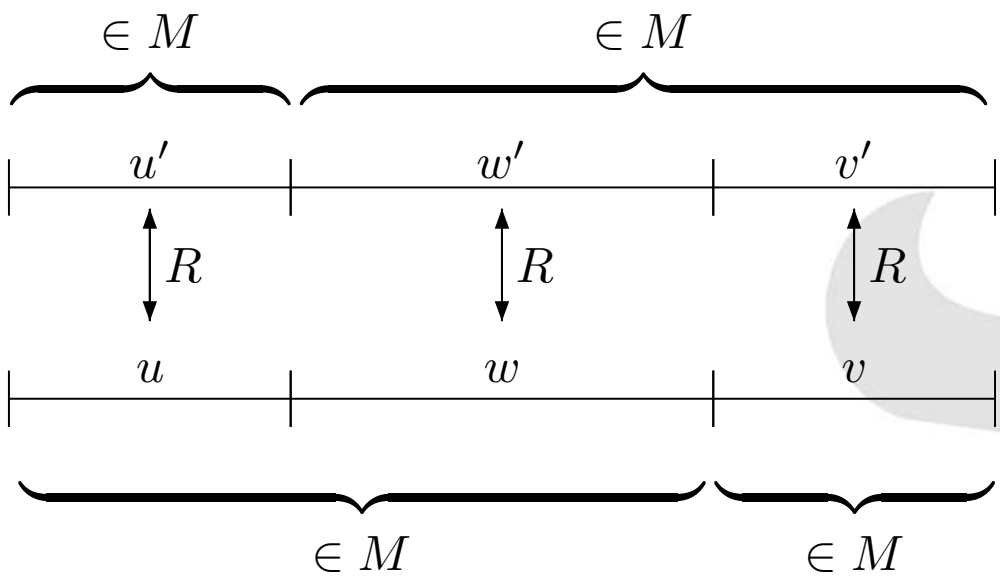
$$(i) \quad M = B^*,$$
$$(ii) \quad B \text{ is an } (R, S)\text{-code.}$$

Theorem 5. X is (R, S)-code $\Leftrightarrow$ X^* is (R, S)-free with minimal generating set X

Theorem 6. M is (R, S)-free $\Leftrightarrow$ M is (R, R)-free and $R_B \subseteq S_B$ for the base B

Stability

A monoid $M \subseteq A^*$ is (R, S)-*stable* if $\forall u, v, w, u', v', w' \in A^*$:



$$\Rightarrow u, w \in M, u \, S \, u'$$

Stability

Theorem 7 (Generalized Schützenberger's criterium).

$$M \text{ is } (R, S)\text{-free} \Leftrightarrow M \text{ is } (R, S)\text{-stable}$$

Stability

Theorem 7 (Generalized Schützenberger's criterium).

$$M \text{ is } (R,S)\text{-free} \Leftrightarrow M \text{ is } (R,S)\text{-stable}$$

Theorem 8 (Generalized Tilson's result). Any nonempty intersection of (R,S)-free monoids of A^* is (R,S)-free.

Hulls

- $\mathcal{F}_{(R,S)}(X) = \{M \mid X^* \subseteq M \subseteq A^*, M \text{ is } (R,S)\text{-free}\}$

Hulls

- $\mathcal{F}_{(R,S)}(X) = \{M \mid X^* \subseteq M \subseteq A^*, M \text{ is } (R,S)\text{-free}\}$
- If $\mathcal{F}_{(R,S)}(X) \neq \emptyset$, then there exists

$$\boxed{F_{(R,S)}(X) = \bigcap_{M \in \mathcal{F}_{(R,S)}(X)} M}$$

Hulls

- $\mathcal{F}_{(R,S)}(X) = \{M \mid X^* \subseteq M \subseteq A^*, M$ is (R,S)-free$\}$
- If $\mathcal{F}_{(R,S)}(X) \neq \emptyset$, then there exists

$$\boxed{F_{(R,S)}(X) = \bigcap_{M \in \mathcal{F}_{(R,S)}(X)} M}$$

- $F_{(R,S)}(X)$ is the (R,S)-*free hull* of X

Hulls

- $\mathcal{F}_{(R,S)}(X) = \{M \mid X^* \subseteq M \subseteq A^*, M$ is (R,S)-free$\}$
- If $\mathcal{F}_{(R,S)}(X) \neq \emptyset$, then there exists

$$\boxed{F_{(R,S)}(X) = \bigcap_{M \in \mathcal{F}_{(R,S)}(X)} M}$$

- $F_{(R,S)}(X)$ is the (R,S)-*free hull* of X
- **Theorem 9.** Let $F_R = F_{(R,R)}(X)$.
  $F_{(R,S)}(X)$ exists $\Leftrightarrow R_{F_R} \subseteq S_{F_R}$. Then $F_{(R,S)}(X) = F_R$.

Hulls

$$C_f(X) = \{(u,v) \in X \times X \mid (u,v) \notin R, uX^* \cap R(vX^*) \neq \emptyset\}.$$

Algorithm 1 (*Base of (R,R)-free hull A_f*). *Input: finite $X \subseteq A^+$. Set $X_0 = X$, and iterate for $j \geq 0$.*

1. *Choose $(u,v) \in C_f(X_j, R)$ such that $u = u'u''$, where $|u'| = |v|$ and $u'' \in A^+$. If no such pair exists, then stop and return $A_f(X) = X_j$.*
2. *Set $R'(u) = \{\mathsf{pref}_{|u'|}(w) \mid w \in (R_{X_j})^+(u)\}$ and set $R''(u) = \{\mathsf{suf}_{|u''|}(w) \mid w \in (R_{X_j})^+(u)\}$, where $(R_{X_j})^+$ is the transitive closure of R_{X_j}.*
3. *Set $X_{j+1} = \big(X_j \setminus (R_{X_j})^+(u)\big) \cup R'(u) \cup R''(u)$.*

Defect effect

Theorem (Defect theorem). Let $X \subseteq A^+$ be a finite set and let B be the base of the free hull of X. Then $|B| \leq |X|$, and the equality holds if and only if X is a code.

Defect effect

Theorem (Defect theorem). Let $X \subseteq A^+$ be a finite set and let B be the base of the free hull of X. Then $|B| \leq |X|$, and the equality holds if and only if X is a code.

- $G_R(X) = (V, E)$: $V = X$, $(u, v) \in E \Leftrightarrow u\,R\,v$
- $c(X, R) =$ the number of connected components of $G_R(X)$.

Defect effect

Theorem (Defect theorem). Let $X \subseteq A^+$ be a finite set and let B be the base of the free hull of X. Then $|B| \leq |X|$, and the equality holds if and only if X is a code.

- $G_R(X) = (V, E)$: $V = X$, $(u, v) \in E \Leftrightarrow u\,R\,v$
- $c(X, R) =$ the number of connected components of $G_R(X)$.

Theorem 11 (Generalized defect theorem). Let X be a finite subset of A^* and let B be the base of the (R, R)-free hull of X. Then $c(B, R) \leq c(X, R)$, and the equality holds if and only if X is an (R, R)-code.

Defect effect

- *pcodes*: $(R_\uparrow, \iota)$-codes over $A_\diamondsuit$

Defect effect

- *pcodes*: $(R_\uparrow, \iota)$-codes over $A_\diamondsuit$
- *pfree*: monoid is generated by a pcode

Defect effect

- *pcodes*: $(R_\uparrow, \iota)$-codes over $A_\diamondsuit$
- *pfree*: monoid is generated by a pcode

Corollary 1 (Defect theorem of partial words). Let X be a finite set of partial words, i.e., a set of words over the alphabet $A_\diamondsuit$. Suppose that pfree hull of X exists and let B be its base. Then $|B| \leq |X|$, and the equality holds if and only if X is a pcode.

References

[1] J. Berstel and L. Boasson, Partial words and a theorem of Fine and Wilf. Theoret. Comput. Sci. 218, 135–141, 1999.

[2] J. Berstel and D. Perrin, Theory of Codes. Academic press, New York, 1985.

[3] J. Berstel, D. Perrin, J.F. Perrot and A. Restivo, Sur le théorème du défaut. J. Algebra 60, 169–180, 1979.

[4] F. Blanchet-Sadri, Codes, orderings, and partial words. Theoret. Comput. Sci. 329, 177–202, 2004.

[5] F. Blanchet-Sadri and M. Moorefield, Pcodes of partial words. Manuscript, 2005.

[6] M. Crochemore and W. Rytter, Jewels of Stringology. World Scientific Publishing, 2002.

References

[7] M.R. Garey and D.S. Johnson, Computer and Intractability: A Guide to the Theory of NP-Completeness. Freeman, New York, 1979.

[8] V. Halava, T. Harju and T. Kärki, Relational codes of words, TUCS Tech. Rep. 767, Turku Centre for Computer Science, Finland, 1–16, April 2006.

[9] V. Halava, T. Harju and T. Kärki, Defect theorems with compatibility relations, TUCS Tech. Rep. 778, Turku Centre for Computer Science, Finland, 1–26, August 2006.

[10] T. Harju and J. Karhumäki, Many aspects of Defect Theorems. Theor. Comput. Sci. 324, 35–54, 2004.

[11] A.A. Sardinas and G.W. Patterson, A necessary and sufficient condition for the unique decomposition of coded messages. IRE Internat. Conv. Rec. 8, 104–108, 1953.

[12] B. Tilson, The intersection of free submonoids of free monoids is free. Semigroup forum 4, 345–350, 1972.