Independent sets, codes and their properties

Meixia TANG [1,a], Yu LIU [2,b]

[1]*School of Information Engineering, Nanning College for Vocational Technology, Nanning, Guangxi*
[2]*Modern Education Technology Center, Nanning College for Vocational Technology,Nanning,Guangxi*
[a]*smiletmx@126.com,* [b]*ly102091002@163.com*

Abstract.

The concept that strict binary relation on free monoids is introduced, some characterizations for strict binary relations is given, the ordering properties of the set of all strict binary relations as well as some subsets of it are exhibited. Moreover, it is proved that the independent languages of co-compatible quasi-strict relations are codes.

Keywords: code; independent set; quasi-strict binary relation; co-compatible binary relation

Introduction

This paper will introduce the concept of quasi-strict relations, some order relations and codes, and discuss the relationship between independent language of quasi-strict relations and code. At last prove a nonempty co-compatible quasi-strict relation independent sets are code.

I. Basic notions and notation

Let X be an alphabet and let X^* be the free monoid generated by X. Any element of X^* is called a word over X and any subset A of X^* is called a language over X. Let $X^+ = X^* - \{1\}$, where 1 is the empty word. We let $\lg(w)$ denote the length of the word w. For any $A, B \subseteq X^*$,

let $AB = \{ab \mid a \in A, b \in B\}$. A non-empty language $A \subseteq X^+$ is called a code if $a_1 a_2 \cdots a_n = b_1 b_2 \cdots b_m, a_i, b_j \in A, i = 1, 2, \cdots, n, j = 1, 2, \cdots, m$, implies $n = m$ and $a_i = b_i$ for $i = 1, 2, \cdots, n$. A code A is said to be a prefix (suffix) code if $A \bigcap AX^+ = \phi$ ($A \bigcap X^+ A = \phi$).

A binary relation ρ on X^* is a subset of $X^* \times X^*$.

A binary relation ρ defined on a set A is called a partial order relation if for all a, b, and c in A, we have that:

(i) a ρ a (reflexivity);

(ii) if a ρ b and b ρ a then a = b (antisymmetry);

(iii) if a ρ b and b ρ c then a ρ c (transitivity).

The set A is called a partially ordered set, poset for short, denoted by (A, ρ).

We call ρ a quasi-strict binary relation on X^* if for all $a, b \in X^*$,

(i) $(a, a) \in \rho$ and $(1, a) \in \rho$

(ii) $(a, b) \in \rho$ and $\lg(a) = \lg(b)$ implies $a = b$

It is clear that ρ is reflexive.

We call ρ a strict binary relation on X^* if for all $a, b \in X^*$,

(i) $(a, a) \in \rho$ and $(1, a) \in \rho$

(ii) $(a, b) \in \rho$ implies $\lg(b) \geq \lg(a)$

(iii) $(a, b) \in \rho$ and $\lg(a) = \lg(b)$ implies $a = b$

It is clear that ρ is reflexive, antisymmetric and transitive.

We let ω_ρ denote the relation $\rho \bigcup \rho^{-1}$ (ρ^{-1} is the Inverse relation of ρ), obviously it is reflexive and symmetric.

Let A_1 be any subset of a set A_2 with a partial ordering $\leq$ on A_2. An element m is said to be a maximal element in A_1 if for every $a \in A_1$, $m \leq a$ implies $a \leq m$. The minimal elements of A_1 is defined correspondingly.

A non-empty subset D of X^* is said to be dependent with respect to a binary relation ρ defined on X^* or simply ρ-dependent if there exist two distinct words u and v in D such that $u\rho v$. As an exception, let $\{1\}$ be ρ-dependent for every ρ. A set $H \subseteq X^*$ is said to be ρ-independent whenever H is not ρ-dependent. The family of all ρ-dependent (ρ-independent) subsets of X^* is called a ρ-dependence (ρ-independence) in X^* and denoted by D_ρ (H_ρ, respectively). Every word $\neq 1$ is in H_ρ for every ρ.

A ρ-dependence relation denoted by $\overset{\rho}{\propto}$ is a binary relation defined on X^* such that $u \overset{\rho}{\propto} v$ if and only if $u\rho v$ or $v\rho u$. The symbol $u \overset{\tilde{\rho}}{\propto} v$ means $u \tilde{\rho} v$ and $v \tilde{\rho} u$.

Let M be a ρ-independent subset of S. Then M is called a maximal ρ-independent subset of S if $M \cup \{x\}$ is ρ-dependent for all $x \in S \setminus M$.

II. Some order relations and codes

We define now the following strict binary relations on X^*:

(i) $\rho_P = \left\{ (u, ux) \mid u \in X^*, x \in X^* \right\}$.

(ii) $\rho_s = \{(u, xu) \mid u \in X^*, x \in X^*\}$.

(iii) $\rho_d = \{(u, y) \mid y = ux$ and $y = wu$ for some $x, w \in X^*\}$.

(iv) $\rho_c = \{(u, y) \mid y = ux = xu$ for some $x \in X^*\}$.

(v) $\rho_e = \{(u, y) \mid u = u_1 u_2 \cdots u_n, y = y_1 u_1 y_2 u_2 \cdots y_n u_n y_{n+1}$, for some $n \geq 0$, and $u_i, y_j \in X^*\}$.

(vi) $\rho_u = \{(u, y) \mid u = y$ or $\lg(u) < \lg(y)\}$.

(vii) $\rho_b = \{(u, y) \mid y = ux$ or $y = wu$ for some $x, w \in X^*\}$.

(viii) $\rho_i = \{(u, y) \mid y \in X^* u X^*\}$.

(ix) $\rho_o = \{(u, y) \mid u = u_1 u_2$ and $y \in u_1 X^* u_2\}$.

In general $\rho_c \subset \rho_d \subset \rho_p \subset \rho_i \subset \rho_e \subset \rho_u$ and $\rho_d \subset \rho_s \subset \rho_i$. It is easy to see that the class of all prefix codes, all hypercodes, all infix codes, all bifix codes, all outfix codes and the class of all suffix codes over X are exactly the class of all independent sets of ρ_p, ρ_e, ρ_i, ρ_b, ρ_o and ρ_s respectively. Obviously, the strict binary relations ρ_p, ρ_s, ρ_e, ρ_d, ρ_u, ρ_b, ρ_i, ρ_o and ρ_c are partial orders on X^*.

PROPOSITION1. Let X be an alphabet and let $x, y \in X^+$, $\lg(y) \geq \lg(x)$. Then the following are equivalent:

(i) $y = ux = xu$ for some $u \in X^*$.

(ii) $xy = yx$.

(iii) $x = w^n$, $y = w^{n+r}$, where $n \geq 1, r \geq 0$ and w is a primitive word over X.

(iv) $\{x, y\}$ is not a code.

PROPOSITION2. If X contains more than one element, then there is no strict binary relation ρ defined on X^* such that the class of all independent sets is exactly the class of all codes over X .

PROOF: Let $X = \{a, b, \cdots\}$ where $a \neq b$. Suppose ρ is a strict binary relation such that the class of all independent sets is exactly the class of all codes. Since every prefix code and every suffix code is a code, we can conclude that $\rho \subseteq \rho_p$ on X^+ and $\rho \subseteq \rho_s$ on X^+ . It follows that for all $u, v \in X^+$, $(u, v) \in \rho$ implies that $v = ux$ and $v = yu$ for some $x, y \in X^*$. The set $A = \{ab^2, ba, ab, b^2a\}$ is not a code, because $ab^2 \cdot ba = ab \cdot b^2a$. However A is an independent set with respect to ρ , a contradiction!

PROPOSITION3. If A is a code over X , then A is an independent set with respect to ρ_c .

PROOF: Let A be a code. The case when A contains only one word is trivial. Now let $u, v \in A$ such that $(u, v) \in \rho_c$, $u \neq v$ .Then by definition $v = ux = xu$ for some $x \in X^+$ .We have $uv = u(xu) = (ux)u = vu$ .This contradicts the fact that A is a code. Hence A is an independent set with respect to ρ_c .

An independent set with respect to ρ_d may not be a code. For example let X = {a, b}. Then $A = \{ab^2, ba, ab, b^2a\}$ is an independent set with respect to ρ_d but A is not a code.

Remark that any code is ρ_c-independent. If every ρ-independent set is a code, then $\rho_c \subseteq \rho$.

III. Co-compatible binary relations on X^*

Let ρ be a binary relation on X^* and let $[\rho] = \overline{\omega_\rho}$, i.e. $[\rho]$ is the complement of ω_ρ, $[\rho]$ is always a symmetric relation. If A is a non-empty set of X^+, then A is ρ-independent if and only if $x[\rho]y$ for every $x, y \in A, x \neq y$.

A binary relation ρ is said to be compatible if

(i) $(x, y) \in \rho$ and $z \in X^*$ imply (xz, yz) and $(zx, zy) \in \rho$.

(ii) (x_1, x_2) and $(y_1, y_2) \in \rho$ imply $(x_1 y_1, x_2 y_2) \in \rho$.

It is well known that if ρ is a reflexive and transitive binary relation, then (i) is equivalent to (ii).

A binary relation ρ is said to be co-compatible if and only if $[\rho]$ is compatible. The strict binary relations ρ_p, ρ_s and ρ_e are co-compatible while ρ_d is not.

PROPOSITION4. Let ρ be a reflexive relation that is co-compatible. Then every ρ-independent set is a code.

PROOF: Since ρ is a reflexive binary relation by assumption, we have $x[\rho]x$ for all $x \in X^*$. Now let $A \subseteq X^+$ be a ρ-independent set. Then $A, A^2, \cdots, A^i, \cdots$ are ρ-independent sets, since ρ is co-compatible.

Suppose A is not a code. Then there exist $x_i, y_j \in A$ such that $x_1 x_2 \cdots x_m = y_1 y_2 \cdots y_n$ for some $m, n \geq 1$ and $x_1 \neq y_1$. We have then $x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n = y_1 y_2 \cdots y_n x_1 x_2 \cdots x_m$.

And $z_1 = x_2 \cdots x_m y_1 y_2 \cdots y_n \neq y_2 \cdots y_n x_1 x_2 \cdots x_m = z_2$ where $z_1, z_2 \in A^{m+n-1}$. Hence $x_1 [\rho] y_1$ and $z_1 [\rho] z_2$ and therefore $x_1 z_1 [\rho] y_1 z_2$ holds. This is a contradiction, since $x [\varnothing] x$ for all $x \in X^*$.

Summary

Code is the most basic tool of information processing, coding theory is the core of formal linguistics, and is often regarded as a separate branch of theoretical computer science and combinatorial mathematics. Prefix code(especially Hoffman code and ASCII code) is the most widely used code. The judge and generation of code is the key problem of coding theory. Notice that some code (such as the prefix, suffix code, super code and infix code etc.)can be defined as a independent language of some relation on free monoids. Introduce a new question: what relation on free monoids that the independent language of it is a code? This paper discusses some properties of independent set and the code, and this problem is solved at last .

References

[1] H.J.Shyr, "Free Monoids and Languages", Lecture Notes, Hon Min Book Company, Taichung, 2001.

[2] S.Ginsburg, "Mathematical Theory of Context-Free Languages", McGraw-Hill, New-York, 1966.

[3] L.H.Haines, on free monoid partially ordered by embedding, J. Combina-torial Theory 6, 94-98, 1969.

[4] H.J.Shyr, and G.Thierren, Hypercodes, Information and Control, Vol. 24, N0.1, 45-54, 1974.

[5] H.K.Hsiao,Y.T.Yeh,S.S.Yu, Dependences related to strict binary relations, Theoretical Computer Science 347, 306-324, 2005.

[6] H.Jurgensen, S.S.Yu, Relations on free monoids, their independent sets, and codes, Internat. J. Comput.Math. 40, 17-46,1991.

[7] J.Berstel and D.Perrin, "Theory of Codes",Academic Press,Version 8 ,2002.

[8] S.S.Yu, "Languages and Codes", Lecture Notes, Tsang Hai Book Publishing Co., Taichung, 2005.

[9] H.J.Shyr and G.Thierrin, Codes and Binary Relations, Lecture Notes 586,180-188, Springer-Verlag, 1976.

[10] H.Jurgensen,S.Konstantinidis, the hierarchy of codes, lecture notes in computer science; Vol.710, 50-68, 1993.

[11] H.Jurgensen, Syntactic Monoids of Codes, report327, Dep. Comp. Sci., The University of Western Ontario, 1992.

[12] H.J.Shyr, S.S.Yu, Intercodes and Some Related Properties, Soochow J. of Math., Volume 16, No.1, 95-107, 1990.

[13] John M.Howie, Automata and Languages, Clarendon Press, 1991.

[14] F. Blanchet-Sadri, Codes, orderings, and partial words, Theoretical Computer Science, Volume 329, Issue 1-3, 177-202, 2004.