About the Triangle Conjecture

Christophe Cordero

Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est Marne-la-Vallée, France

Introduction to Coding Theory

Imagine a communication between a source $\mathcal{S}$ and a target $\mathcal{T}$, where they first agree on a table to encode and decode each character. This table can be a fixed-length code

$$\mathcal{S} \xrightarrow[\cdots|0010|1001|0010|0001]{} \mathcal{T}$$

Table	
0001	a
0010	b
⋮	

or it can be a variable-length code

$$\mathcal{S} \xrightarrow[\cdots|11|10001|01|01|11]{} \mathcal{T}$$

Table	
11	a
10001	b
⋮	

This second type of communication works if and only if the messages sent by $\mathcal{S}$ have a **unique decomposition** in the element of the table. This notion is formalised by the notion of code in coding theory.

Definition

A set $X \subset \mathcal{A}^*$ is a **code** if and only if for all $\omega \in X^*$ there exist a unique $n \geq 0$ and a unique sequence $x_1, \ldots, x_n \in X$ such that

$$\omega = x_1 x_2 \cdots x_n.$$

For example, the set $\{aabb, abaaa, b, ba\}$ is not a code because

$$babaaabb = (b)(abaaa)(b)(b) = (ba)(ba)(aabb).$$

A subset $X \subset \mathcal{A}^*$ is **prefix** if no element of X is a proper prefix of another element in X. For example, the set

$$\{b, ab, a^2b, a^3b, a^4b, \ldots\}$$

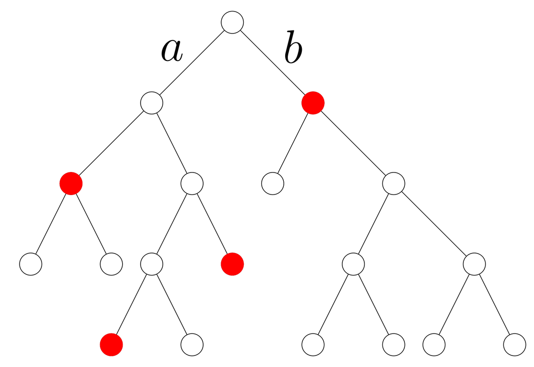
is prefix.

Proposition

Any prefix set different than $\{\varepsilon\}$ is a code.

Why are we interested in prefix code?

1) Because there are easy to produce! For example, the tree



produces the prefix code $\{aa, abaa, abb, b\}$.

2) Because they are easy to decode!
For example, if $\mathcal{T}$ receives the message $aaaaabaaabbabaab$, then $\mathcal{T}$ cuts one by one the prefix of the message that are elements of the code. Thus, the message is decoded as

$$aa, aa, abaa, abb, abaa, b.$$

3) Prefix codes appear in one of the main conjectures in coding theory.

Conjectures

A set $X \subset \mathcal{A}^*$ is **commutatively prefix** if there exists a prefix code P such that

$$\sum_{x \in X} y^{|x|_a} z^{|x|_b} = \sum_{p \in P} y^{|p|_a} z^{|p|_b}.$$

For example, the set $\{a, ba, aabb, baabb, ababb\}$ is commutatively prefix, because it is equivalent to the prefix code $\{a, ba, bbaa, bbaba, bbbaa\}$.

Conjecture (Perrin and Schützenberger)

All finite maximal codes are commutatively prefix.

We study here this conjecture on the particular case of bayonet codes.

A **bayonet** code X is a code such that $X \subset a^*ba^*$. For example, the set $\{ab, abaa, aaaab\}$ is a bayonet code.

Triangle Conjecture (Perrin and Schützenberger)

A finite bayonet code is either commutatively prefix or it is not included in a finite maximal code.

It is easy to determine if a bayonet code is commutatively prefix.

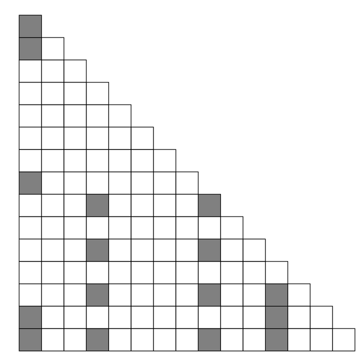
Proposition

A bayonet code X is commutatively prefix if and only if

$$\left|X \cap \mathcal{A}^{\leq n}\right| \leq n, \text{ for all } n \geq 0.$$

In 1984, Shor found the bayonet code

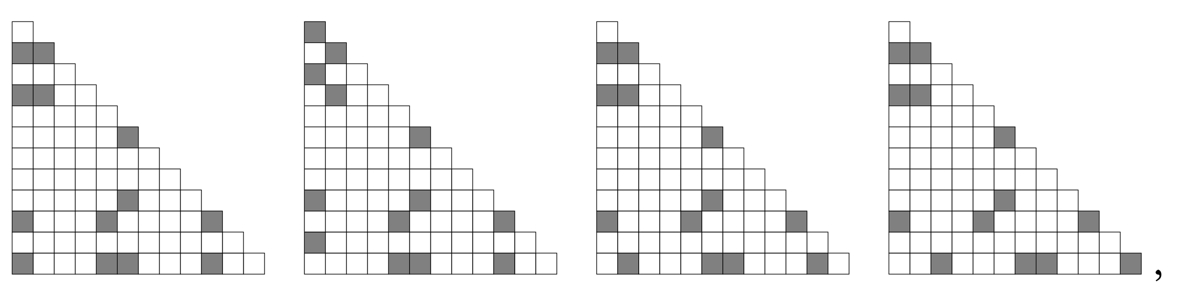
$$\begin{Bmatrix} b, & ba, & ba^7, & ba^{13}, & ba^{14}, \\ a^3b, & a^3ba^2, & a^3ba^4, & a^3ba^6, & \\ a^8b, & a^8ba^2, & a^8ba^4, & a^8ba^6, & \\ a^{11}b, & a^{11}ba, & a^{11}ba^2 & & \end{Bmatrix} \longrightarrow$$



with 16 elements that is included in $\mathcal{A}^{\leq 15}$. Hence, it is a non-commutatively prefix code. We do not know if it is included in a finite maximal code but it is the only non-commutatively prefix code that was known. In here, we show the results of our computer exploration in order to find new non-commutatively prefix code.

Results of our Computer Exploration

For $n \leq 11$: 0 code, $n = 12$:



$n = 13, 14$: 0 code, $n = 15$: 76 codes, $n = 16$: at least 50 codes, $n = 17$: at least 6 codes...

We improved a partial answer to the following question:

Question from Shor

What is the maximum value of $\frac{|X|}{n}$ where X is a code belonging to $a^*ba^* \cap \mathcal{A}^{\leq n}$ and n an integer?

Partial answer (from Shor, Hansel, and us): this value is between $\frac{16}{15} \leq \frac{13}{12}$ and $1 + \frac{1}{\sqrt{2}}$.

Factorisations of Cyclic Groups

Proposition and definition

For all finite maximal code X and for any letter $x \in \mathcal{A}$, there exists an integer k such that $x^k \in X$. Such an integer is called the **order** of the letter x.

This notion is linked to the factorisation theory. Given $n \geq 1$, the ordered pair $(L, R) \subset [0, n[^2$ is a **factorisation** of $\mathbb{Z}/n\mathbb{Z}$ if

$$\forall k \in [0, n[, \exists!(\ell, r) \in L \times R \text{ such that } k = \ell + r \bmod n.$$

For example, the ordered pair $(\{1, 3, 5\}, \{1, 2, 7, 8\})$ is a factorisation of $\mathbb{Z}/12\mathbb{Z}$. The following theorem shows the link between factorisation theory and coding theory.

Theorem (Restivo, Salemi, and Sportelli)

If X is a finite maximal code such that $b, a^n \in X$ then (L, R) is a factorisation of $\mathbb{Z}/n\mathbb{Z}$, where

$$L := \left\{k \bmod n : a^k b^+ \in X\right\} \text{ and } R := \left\{k \bmod n : b^+ a^k \in X\right\}.$$

In the next section, we use the contraposition of the following theorem.

Theorem (Sands)

If (L, R) is a factorisation of $\mathbb{Z}/n\mathbb{Z}$ and p is an integer relatively prime to $|L|$ then (pL, R) is a factorisation of $\mathbb{Z}/n\mathbb{Z}$.

Applications of Factorisations of Cyclic Groups

We recall that Shor's code is

$$\begin{Bmatrix} b, & ba, & ba^7, & ba^{13}, & ba^{14}, \\ a^3b, & a^3ba^2, & a^3ba^4, & a^3ba^6, & \\ a^8b, & a^8ba^2, & a^8ba^4, & a^8ba^6, & \\ a^{11}b, & a^{11}ba, & a^{11}ba^2 & & \end{Bmatrix}$$

Suppose that it is a counter-example to the triangle conjecture. Then there exists a factorisation of the form

$$(L \supseteq \{0, 3, 8, 11\}, R \supseteq \{0, 1, 7, 13, 14\}).$$

However, we do not know any of these factorisations! Note that $(L, 3R)$, $(L, 5R)$, $(L, 8R)$, and $(L, 11R)$ are not factorisations. Thus, thanks to Sands Theorem, we know that $3|n$, $5|n$, $2|n$, and $11|n$. Hence n is a multiple of

$$2 \times 3 \times 5 \times 11 = 330.$$

Let us apply the same strategy to the codes we have found.

We found 27 non-commutatively prefix codes containing b. For 7 of them we found some factorisations. For example, the factorisation associated to the code

$$\begin{Bmatrix} b, & ba^2, & ba^8, & ba^{10}, \\ aba^8, & aba^{10}, & & \\ a^4b, & a^4ba^2, & & \\ a^5b, & a^5ba^3, & a^5ba^6, & \\ a^9b, & a^9ba^2 & & \end{Bmatrix}$$

has of the form

$$(L \supseteq \{0, 4, 5, 9\}, R \supseteq \{0, 2, 8, 10\}).$$

We found the following infinite set of such factorisations: for $n \geq 2$,

$$\left(\{0, 4, 5, 9\}, \bigsqcup_{0 \leq i < n} \{8i, 8i + 2\}\right)$$

is a factorisation of $\mathbb{Z}/8n\mathbb{Z}$.

For the other 20 codes containing b, we did not found factorisations associated to them but we computed the following lower bound on the order of the letter a:

Nb of codes	Order of the letter a
2	$2 \times 3 \times 5 \times k = 30k$, with $k \geq 3$
6	$2 \times 3 \times 11 \times k = 66k$, with $k \geq 3$
2	$2 \times 3 \times 5 \times 11 \times k = 330k$, with $k \geq 4$
4	$2 \times 3 \times 5 \times 13 \times k = 390k$, with $k \geq 4$
4	$2 \times 3 \times 5 \times 13 \times k = 390k$, with $k \geq 3$
2	$2 \times 5 \times 13 \times k = 130k$, with $k \geq 3$

Journées Nationales du *GDR Informatique Mathématique* du 11 au 14 mars 2019 à Orléans en France.