1 Keywords

distinct 0

Theorem 1.1. *Any two ~~unique~~ keywords w and x s.t. $w = w_1 w_2 ... w_n$, $x = x_1 x_2 ... x_n$ in a binary word list must be of the form $w_i = x_i$; $i \in \{2, 3, ..., n\}$.*

1

Proof. Let A be a word list with unique keywords w, x, s.t. $w = w_1 w_2 ... w_n$, and $x = x_1 x_2 ... x_n$ where $w_i, x_i, y_i \in \{0, 1\}$.

By definition of keyword, $\forall a \in A$, $\forall a' \subset a$, $\forall w' \subset we$ s.t. $e \in \{0,1\}$, if a' and w' are the same length, then $\Big| |a'| - |w'| \Big| \leq 1$. This property is the same for x. Therefore, the following words are all contained in A.

can this be more clear.

$$w_2 w_3 ... w_n 0$$
$$w_2 w_3 ... w_n 1$$
$$x_2 x_3 ... x_n 0$$
$$x_2 x_3 ... x_n 1$$

Note that if $w_n \neq x_n$ then we get either case

$$\begin{array}{lcl} w_2 w_3 ... w_{n-1} 00 & & w_2 w_3 ... w_{n-1} 10 \\ w_2 w_3 ... w_{n-1} 01 & & w_2 w_3 ... w_{n-1} 11 \\ x_2 x_3 ... x_{n-1} 10 & \textit{or} & x_2 x_3 ... x_{n-1} 00 \\ x_2 x_3 ... x_{n-1} 11 & & x_2 x_3 ... x_{n-1} 01 \end{array}$$

In either case, w and x would violate the definition of ~~a keyword~~ balanced because $|11| - |00| = 2 > 1$. Therefore $w_n = x_n$.

Assume that

mention proof by induction

$$w_n = x_n$$
$$w_{n-1} = x_{n-1}$$
$$\vdots$$
$$w_{n-k} = x_{n-k}$$

For convenience, let $e_i = w_{n-i}$; $i \in \{2, ..., k\}$. Note that if $w_{n-k-1} \neq x_{n-k-1}$ then we get either case

$$\begin{array}{lcl} w_2 w_3 ... w_{n-k-2} 0 e_i 0 & & w_2 w_3 ... w_{n-k-2} 1 e_i 0 \\ w_2 w_3 ... w_{n-k-2} 0 e_i 1 & & w_2 w_3 ... w_{n-k-2} 1 e_i 1 \\ x_2 x_3 ... x_{n-k-2} 1 e_i 0 & \textit{or} & x_2 x_3 ... x_{n-k-2} 0 e_i 0 \\ x_2 x_3 ... x_{n-k-2} 1 e_i 1 & & x_2 x_3 ... x_{n-k-2} 0 e_i 1 \end{array}$$

In either case, w and x would violate the definition of a keyword because $|1 e_i 1| - |0 e_i 0| = 2 > 1$. Therefore $w_{n-k} = x_{n-k}$, and by induction

$$w_i = x_i; \;\; i \in \{2, 3, ..., n\}$$

.

$\square$

Lemma 1.2. *Any word list has at most two keyword*s.

Proof. Let A be a word list of keywords w, x and y s.t. $w = w_1 w_2 ... w_n$, $x = x_1 x_2 ... x_n$ and $y = y_1 y_2 ... y_n$. Then by theorem 1.1

$$x_i = w_i = y_i; \ i \in \{2, 3, ..., n\}$$

Since this is a binary language, there are only two possible values for the first letter. Hence, it is impossible for w_1, x_1, and y_1 to all be unique. At least two of the letters must be the same.

Therefore $w = x$ or $x = y$ or $y = w$. Therefore A must have at most two keywords. □

Lemma 1.3. *If a word list has two unique keywords w, and x s.t. $w = w_1 w_2 ... w_n$, $x = x_1 x_2 ... x_n$, then the keywords are of the form $0w_2 w_3 ... w_n$ and $1w_2 w_3 ... w_n$.* ~~*Hence*~~ T*the largest balanced subsets of $\{w0, w1, x0, x1\}$ are $\{w0, w1, x0\}$ and $\{w1, x0, x1\}$ where $w = 0w_2 w_3 ... w_n$ and $x = 1w_2 w_3 ... w_n$.*

Given 2 distinct keywords X, w

Proof. Let A be a word list with two unique keywords w, and x s.t. $w = w_1 w_2 ... w_n$,
$x = x_1 x_2 ... x_n$. By theorem 1.1

$$x_i = w_i = y_i; \ i \in \{2, 3, ..., n\}$$

Therefore, to maintain uniqueness, $w_1 \neq x_1$. Becuase this is a binary language, either $w_1 = 0, x_1 = 1$ or $w_1 = 1, x_1 = 0$. □

This leads splitting

Theorem 1.4. *If a word list has two keywords $0w$ and $1w$ then w is a palindrome.*

Proof. Let $0w$ and $1w$ be two keywords of a word list, A. Then by theorem 1.1, ? the following words are in A.

by def'n of keyword

$$\left\{ \begin{array}{c} 0w \\ 1w \\ w0 \\ w1 \end{array} \right\}$$

mistake

Let's represent the characters of w as $w = w_0 w_1 ... w_n$. For sake of contradiction, assume that $w_1 \neq w_n$. Then either case of words would be in A

In either case the set of words in A would be ... OR ...

where did the cases come from?

$$\left\{ \begin{array}{c} 00w_1 ... w_{n-1}1 \\ 10w_1 ... w_{n-1}1 \\ 0w_1 ... w_{n-1}10 \\ 0w_1 ... w_{n-1}11 \end{array} \right\} \qquad \left\{ \begin{array}{c} 01w_1 ... w_{n-1}0 \\ 11w_1 ... w_{n-1}1 \\ 1w_1 ... w_{n-1}00 \\ 1w_1 ... w_{n-1}01 \end{array} \right\}$$

Cases are $w_1 = 0$, $w_n = 1$ $\qquad$ $w_1 = 1$, $w_n = 0$

spell this all out, this ___ and this ___ are not balanced but in A.

$\rightarrow\leftarrow$ contradiction

In either case, there exist subwords 00 and 11 which would result in an unbalanced word list. Therefore $w_1 = w_n$. ¶ Assume for some k that $w_i = w_{n-i}$; $0 \leq i \leq k$. For sake of contradiction, also assume that $w_{i+1} \neq w_{n-i}$. Then either case of words would be in A.

what are the cases

$w_i = w_{n-i+1}$

adjust subscripts

$0w_0...w_i0...1w_{n-i}..w_n$	$0w_0...w_i1...0w_{n-i}..w_n$
$1w_0...w_i0...1w_{n-i}..w_n$	$1w_0...w_i1...0w_{n-i}..w_n$
$w_0...w_i0...1w_{n-i}..w_n0$	$w_0...w_i1...0w_{n-i}..w_n0$
$w_0...w_i0...1w_{n-i}..w_n1$	$w_0...w_i1...0w_{n-i}..w_n1$

name the sets

In the left case $0w_0...w_i0$ is out of balance with $1w_{n-i}..w_n1$. In the right case $1w_0...w_i1$ is out of balance with $0w_{n-i}..w_n0$. Hence, $w_{n-k-1} = w_{k+1}$. Therefore, through induction, $w = \bar{w}$. $\square$

the # of gen. words is $n+2$.

Lemma 1.5. *Any word list can have at most two children, implicating the tree of word lists is binary.*

Proof. Let A be a word list of n words, two of which are keywords, and $\forall a \in A$ the length of a is $n-1$. Then $\left|\bar{C}(A)\right| = n+2$, with $\forall a \in \bar{C}(A)$ the length of a is n. Hence $\bar{C}(A)$ can only produce two ~~unique~~ lists with ~~complexity~~ size $n+1$.

If A instead has only one keyword. Then $\left|\bar{C}(A)\right| = n+1$. Hence $\bar{C}(A)$ can only produce one unique list of ~~complexity~~ size $n+1$. □

* distinct balanced lists