17-750
1/29/10

Codes & Uniquely Decipherable Codes

Def:

Alphabet $\equiv \Sigma = \{0, \sim, r-1\}$ this lecture $\{0,1\}$

Code Word $\equiv$ string $w = a_1 \cdots a_\ell$ $a_i \in \Sigma$

Code $\equiv$ set of codeword. $\{w_1, \cdots, w_k\}$

Message $\equiv$ string of code words (concatenation)

eg Code $\equiv \{\overset{w_1}{01}, \overset{w_2}{0}, \overset{w_3}{10}\}$

$01010 = w_1 w_2 w_3$ Not Uniquely Decipherable.

$w_2 w_3 w_3 = w_1 w_2 w_2$

Def Code C is UD if every message is UD.

If $p \cdot s = w \in \Sigma^*$ then p prefix of w & s is suffix of w.

Testing UD

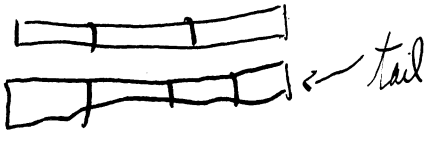
Input: Code $= \{c_1, \ldots, c_K\}$

Def A non-empty word (string) t is a tail if $\exists$ messages $c_1 \cdots c_m$ & $c_1' \cdots c_n'$ s.t.

1) c_i, c_j' are code words & $c_1 \neq c_1'$
2) t is a suffix of c_n'
3) $c_1 \cdots c_m t = c_1' \cdots c_n'$

Lemma C is UD iff no tail is a code word.

($\Leftarrow$) (contrapositive) if not UD $\Rightarrow \exists$ tail as code word



($\Rightarrow$)($\exists$ tail codeword $\Rightarrow$ not UD)

Same argument

Alg (Generate All Tails)

Generate-All-Tails (tails = $\emptyset$)

1) $\forall C_i \& C_j$ $(i \neq j)$

   a) $C_i = C_j$ return not UD

   b) $\exists s$ s.t. $C_i s = C_j$ or $C_i = C_j s$

      tails = tails $\cup \{s\}$

2) $\forall c \in C \forall t \in$ tails (memoize)

   a) $t = C$ return not UD

   b) $\exists s$ $\overset{2b'}{ts} = C$ or $\overset{2b''}{Cs} = t$

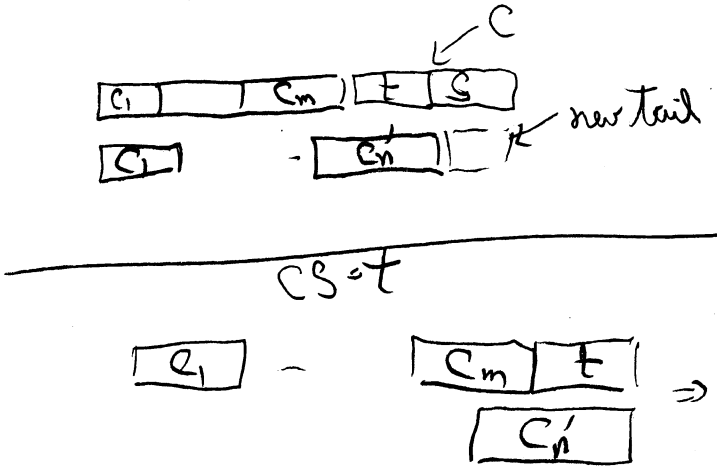
      tails = tails $\cup \{s\}$

3) return C is UD

Claim 1 Alg generates only tails

1b) is OK

2b) $ts = c$



Claim2 Either Alg returns not UD or it generates all Tails

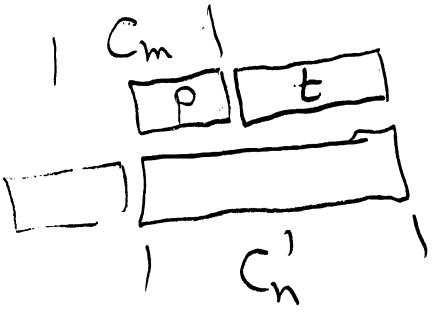
Let t be a tail $c_1 \cdots c_m t$ with min n+m
$c_1' \cdots c_n'$

Def $m(t) = \min\{n+m\}$

Induct on $m(t)$

$m(t) = 1$ done by step 1)

Assume true for $k < n+m$

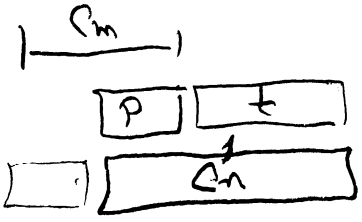


Case a

$P = C_m$ then P is found' trivially by induct an alg returns "not UD"

Case b

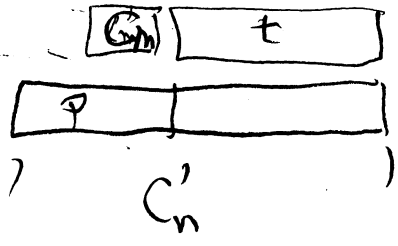
P suffix of C_m



induct $P \in Tails \overset{2b'}{\Rightarrow} t_i \in Tails$

Case c

C_m suffix of P



$C_m t \in Tails \overset{2b''}{\Rightarrow} t \in Tails$

Timing: n words max length l

tails $\leq n \cdot l$

Step 1 $O(n^2 \cdot l)$ $O(n \cdot l)$?

Step 2 # words $= n$

tail $= n \cdot l$

cost per pair l

$O(n^2 \cdot l^2)$