IJMMS 30:8 (2002) 491–504
PII. S0161171202011729
http://ijmms.hindawi.com

NOTE ON DECIPHERABILITY OF THREE-WORD CODES

F. BLANCHET-SADRI and T. HOWELL

Received 29 January 2001

The theory of uniquely decipherable (UD) codes has been widely developed in connection with automata theory, combinatorics on words, formal languages, and monoid theory. Recently, the concepts of multiset decipherable (MSD) and set decipherable (SD) codes were developed to handle some special problems in the transmission of information. Unique decipherability is a vital requirement in a wide range of coding applications where distinct sequences of code words carry different information. However, in several applications, it is necessary or desirable to communicate a description of a sequence of events where the information of interest is the set of possible events, including multiplicity, but where the order of occurrences is irrelevant. Suitable codes for these communication purposes need not possess the UD property, but the weaker MSD property. In other applications, the information of interest may be the presence or absence of possible events. The SD property is adequate for such codes. Lempel (1986) showed that the UD and MSD properties coincide for two-word codes and conjectured that every three-word MSD code is a UD code. Guzmán (1995) showed that the UD, MSD, and SD properties coincide for two-word codes and conjectured that these properties coincide for three-word codes. In an earlier paper (2001), Blanchet-Sadri answered both conjectures positively for all three-word codes $\{c_1,c_2,c_3\}$ satisfying $|c_1| = |c_2| \le |c_3|$. In this note, we answer both conjectures positively for other special three-word codes. Our procedures are based on techniques related to dominoes.

2000 Mathematics Subject Classification: 94A15, 05A99, 94B35.

1. Introduction. Let A be a nonempty finite set or an alphabet; A^* denotes the set of all sequences of finite length (greater than or equal to 0) of elements of A (such sequences are called words on A). The unique sequence of length 0, denoted by ϵ, is called the empty word. *A code* C on A is a nonempty finite subset of $A^+ = A^* \setminus \{\epsilon\}$. The words in C are called code words. *A message* on C is a word in A^* that is a concatenation of code words. The sequence of these code words is a *decoding* or *factorization* of the message. The code C is called

- *uniquely decipherable*, if every message on C has a unique factorization into code words;
- *multiset decipherable*, if any two factorizations of the same message on C yield the same multiset of code words;
- *set decipherable*, if any two factorizations of the same message on C yield the same set of code words.

Every UD code is MSD, and every MSD code is SD. It has been shown that these relationships are proper. The code $C = \{0,0111110,10101,1111\}$ on $\{0,1\}$ is an example of a *proper* MSD code (i.e., an MSD code that is not UD). In fact, the message

$$(0111110)(10101)(1111)(0) = (0)(1111)(10101)(0111110) \tag{1.1}$$

on C has two distinct factorizations into code words [3, 5]. The code

$$C = \{0, 010, 11011, 101101\} \tag{1.2}$$

on $\{0,1\}$ is an example of a *proper* SD code (i.e., an SD code that is not MSD). The message

$$(0)(101101)(11011)(0)(11011)(010) = (010)(11011)(101101)(101101)(0) \tag{1.3}$$

on C has two distinct factorizations with distinct multisets of code words [3, 5]. Guzmán presents a complete list of proper MSD and proper SD four-word codes on $\{0,1\}$ with code words of length less than or equal to 7 [3]. It is decidable whether or not a code C is UD [1, 5, 6, 9] (resp., MSD [5]).

For two-word codes, the UD, MSD, and SD properties coincide [3, 7]. For three-word codes, it is an open question whether or not they coincide. Lempel [7] conjectured that every three-word MSD code is a UD code, and Guzmán [3] conjectured that the UD, MSD, and SD properties coincide for three-word codes. We answered both conjectures positively for all three-word codes $\{c_1, c_2, c_3\}$ satisfying $|c_1| = |c_2| \le |c_3|$ [2]. In this note, we give (in Section 1.1) a brief overview of Head's and Weber's domino technique [5], and give (in Section 2) an application of this approach by proving that Lempel's and Guzmán's conjectures are true for some special three-word codes.

1.1. A domino technique. Let A be an alphabet and C a code on A. The set of all prefixes of words in C will be denoted by $\mathrm{Prefix}(C)$. The *domino graph* associated with C is the directed graph $G = (V, E)$ where

$$V = \left\{ \text{open}, \text{close}, \binom{u}{\epsilon}, \binom{\epsilon}{u} \;\middle|\; u \in \mathrm{Prefix}(C) \setminus \{\epsilon\} \right\}, \tag{1.4}$$

and $E = E_1 \cup E_2 \cup E_3 \cup E_4$ where

$$\begin{aligned}
E_1 &= \left\{ \left(\text{open}, \binom{u}{\epsilon} \right), \left(\text{open}, \binom{\epsilon}{u} \right) \;\middle|\; u \in C \right\}, \\
E_2 &= \left\{ \left(\binom{u}{\epsilon}, \text{close} \right), \left(\binom{\epsilon}{u}, \text{close} \right) \;\middle|\; u \in C \right\}, \\
E_3 &= \left\{ \left(\binom{u}{\epsilon}, \binom{uv}{\epsilon} \right), \left(\binom{\epsilon}{u}, \binom{\epsilon}{uv} \right) \;\middle|\; v \in C \right\}, \\
E_4 &= \left\{ \left(\binom{u}{\epsilon}, \binom{\epsilon}{v} \right), \left(\binom{\epsilon}{u}, \binom{v}{\epsilon} \right) \;\middle|\; uv \in C \right\}.
\end{aligned} \tag{1.5}$$

The *domino function* associated with C is the mapping d from E to $\left\{ \binom{u}{\epsilon}, \binom{\epsilon}{u} \mid u \in C \right\}$ defined on

- E_1 by $\left(\text{open}, \binom{u}{\epsilon}\right) \mapsto \binom{\epsilon}{u}$ and $\left(\text{open}, \binom{\epsilon}{u}\right) \mapsto \binom{u}{\epsilon}$,
- E_2 by $\left(\binom{u}{\epsilon}, \text{close}\right) \mapsto \binom{u}{\epsilon}$ and $\left(\binom{\epsilon}{u}, \text{close}\right) \mapsto \binom{\epsilon}{u}$,
- E_3 by $\left(\binom{u}{\epsilon}, \binom{uv}{\epsilon}\right) \mapsto \binom{\epsilon}{v}$ and $\left(\binom{\epsilon}{u}, \binom{\epsilon}{uv}\right) \mapsto \binom{v}{\epsilon}$,
- E_4 by $\left(\binom{u}{\epsilon}, \binom{\epsilon}{v}\right) \mapsto \binom{uv}{\epsilon}$ and $\left(\binom{\epsilon}{u}, \binom{v}{\epsilon}\right) \mapsto \binom{\epsilon}{uv}$.

The *domino* associated with an edge e of E is the domino $d(e) = \binom{d_1(e)}{d_2(e)}$. The function d induces mappings d_1 and d_2 from E to $C \cup \{\epsilon\}$ also called domino functions. If $p = e_1 \cdots e_m$ is a path in G, the word $d(e_1) \cdots d(e_m)$ (resp., $d_1(e_1) \cdots d_1(e_m)$, $d_2(e_1) \cdots d_2(e_m)$) will be denoted by $d(p)$ (resp., $d_1(p)$, $d_2(p)$).

A path p in G from *open* to some vertex $\binom{u}{\epsilon}$, respectively, $\binom{\epsilon}{u}$, is trying to find two factorizations of the same message on C into code words beginning with distinct code words. The decodings obtained so far are $d_1(p)$ and $d_2(p)$. The word u in A^* denotes the backlog of the first (resp., second) decoding as against the second (resp., first) one.

Guzmán [4] suggested to look at the *simplified domino graph* and the *domino function* of C. The simplified domino graph of C is a subgraph of the domino graph of C. Replace E_1 by $E_1' = \left\{\left(\text{open}, \binom{\epsilon}{u}\right) \mid u \in C\right\}$, E_2 by $E_2' = \left\{\left(\binom{u}{\epsilon}, \text{close}\right) \mid u \in C\right\}$, V by V' which consists of open, close, and those vertices v in V such that there is a path from open to close that goes through v, and E by E' which consists of those edges e in E such that there is a path from open to close going through e. The simplified domino graph of C is denoted by $G(C)$.

The UD, MSD, and SD properties of a code C can be characterized in terms of its simplified domino graph $G(C)$ and functions d_1 and d_2,

- C is not UD if and only if there is a path in $G(C)$ from open to close [6].
- C is not MSD if and only if there is a path p in $G(C)$ from open to close such that $d_1(p)$ and $d_2(p)$ do not have the same multiset of code words [5].
- C is not SD if and only if there is a path p in $G(C)$ from open to close such that $d_1(p)$ and $d_2(p)$ do not have the same set of code words [3].

As an example, we consider the code $c = \{c_1, c_2, c_3, c_4\}$ on $\{0,1\}$ discussed in [5, 4] ($c_1 = 0, c_2 = 0111110, c_3 = 10101$, and $c_4 = 1111$). The simplified domino graph and function associated with this code are shown in Figure 1.1.

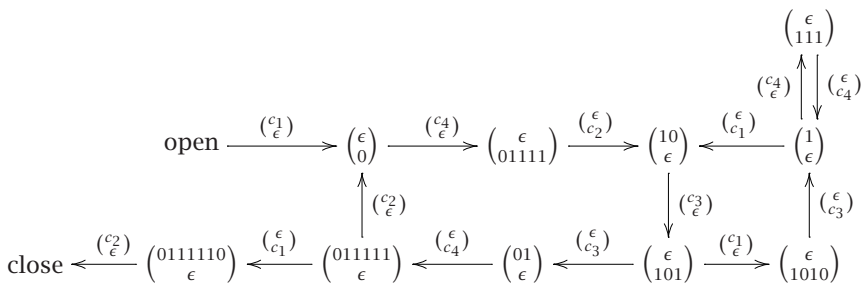

FIGURE 1.1

Each edge e is labeled by $d(e)$. The path

$$p = \text{open}, \binom{\epsilon}{0}, \binom{\epsilon}{01111}, \binom{10}{\epsilon}, \binom{\epsilon}{101}, \binom{01}{\epsilon}, \binom{011111}{\epsilon}, \binom{0111110}{\epsilon}, \text{close} \tag{1.6}$$

is from open to close showing that C is not UD. However, every path p from open to close is such that $d_1(p)$ and $d_2(p)$ have the same multiset of code words showing that C is MSD.

1.2. Preliminary lemmas. We give some preliminary lemmas that are used in Section 2 to prove our main results.

Definition 1.1 (see Lothaire [8]). A nonempty word u on an alphabet A is called *primitive*, if $u = v^n$ for some nonempty word v on A implies that $n = 1$.

Lemma 1.2 (see Lothaire [8]). *Let u be a nonempty word on an alphabet A. There exist a unique primitive word v and a unique positive integer n such that $u = v^n$ (v, denoted by $\sqrt{u}$, is called the root of u and n, denoted by $\exp(u)$, is called the exponent of u). Moreover, all positive powers of u have the same root.*

Lemma 1.3 (see Lothaire [8]). *Let $C = \{u, v\}$ be a two-word on an alphabet A. The* UD, MSD, *and* SD *properties are equivalent to the following properties:*
(1) $uv \neq vu$.
(2) $\sqrt{u} \neq \sqrt{v}$.

Lemma 1.4 (see Blanchet-Sadri [2]). *Let u, x be nonempty words on an alphabet A and let k be a positive integer. If $\{ux, (ux)^k u\}$ is* SD, *then $ux \neq xu$.*

Lemma 1.5. *Let u, v, x, y be nonempty words on an alphabet A satisfying $u \neq v$.*
(1) *The equalities $vyx = yxu = xuy$ cannot hold simultaneously.*
(2) *The equalities $xvy = vyx = yxu$ cannot hold simultaneously.*

Proof. We prove assertion (1) (assertion (2) is proved similarly). Since $yxu = xuy$, by Lemmas 1.2 and 1.3 there exist a unique primitive word w and unique positive integers k and k' such that $y = w^k$ and $xu = w^{k'}$. Put $x = w^\ell z$ and $u = z'w^{\ell'}$ where w is not a prefix of z and w is not a suffix of z'. If $z, z' = \epsilon$, then $u = v$, a contradiction. Otherwise, $z, z' \neq \epsilon$ and the equality $vyx = xuy$ implies that $w = zz' = z'z$. Lemmas 1.2 and 1.3 imply the existence of a unique primitive word w' and unique positive integers m and m' satisfying $z = (w')^m$ and $z' = (w')^{m'}$. We get $w = (w')^{m+m'}$, a contradiction with the fact that w is primitive. □

Lemma 1.6. *Let u, v, x, y be nonempty words on an alphabet A satisfying $u \neq v$. If the equalities $vyx = yxv = yux$ hold simultaneously, then $u = (z'z)^k$, $v = (zz')^k$, $x = (z'z)^\ell z'$, and $y = (zz')^m z$ for some $k > 0$, $\ell \geq 0$, $m \geq 0$, and words z, z' on A.*

Proof. Since $vyx = yxv$, by Lemmas 1.2 and 1.3 there exist a unique primitive word w and unique positive integers k and k' such that $v = w^k$ and $yx = w^{k'}$. Put $y = w^\ell z$ and $x = z'w^{\ell'}$ where w is not a prefix of z and w is not a suffix of z'. If $z, z' = \epsilon$, then $u = v$, a contradiction. Otherwise, $z, z' \neq \epsilon$ and $w = zz'$, and the equality $yxv = yux$ implies that $u = (z'z)^k$ and the result follows. □

Lemma 1.7 (see Blanchet-Sadri [2]). *Let u, v, x be nonempty words on an alphabet A satisfying $u \neq v$. The equalities $ux = xv$ and $vx = xu$ cannot hold simultaneously.*

2. Some special three-word codes. We show that the UD, MSD, and SD properties are equivalent to the special three-word codes described in Theorems 2.1, 2.2, and 2.3.

Let $C = \{c_1, c_2, c_3\}$ be a three-word SD code on alphabet A satisfying $|c_1| < |c_2| < |c_3|$. Let c_2 and c_3 be written, respectively, as $c_1^k u$ and $c_1^k v$ where c_1 is not a prefix

of u, c_1 is not a prefix of v, u is not a prefix of v, $k > 0$, and $u, v \neq \epsilon$. For each of our theorems, we need only to show that C is UD. It is easy to see that $\{c_1, c_2\}$, $\{c_1, c_3\}$, and $\{c_2, c_3\}$ are two-word SD codes, and are therefore MSD and UD codes by Lemma 1.3. When constructing $G(C)$, we see that E_1' consists of the edges $\left(\text{open}, \binom{\epsilon}{c_i}\right)$ and E_2' consists of the edges $\left(\binom{c_i}{\epsilon}, \text{close}\right)$. No E_3- nor E_4-edge leaves $\binom{\epsilon}{c_2}$ and no E_3- nor E_4-edge leaves $\binom{\epsilon}{c_3}$.

Let u_i (resp., v_i, w_i) be the prefix of length i of c_1 (resp., c_2, c_3). We write $c_1 = u_i x_i$. When trying to build a path from open to close, a list of vertices gets generated. The vertices are among $\binom{u_i}{\epsilon}$, $\binom{v_i}{\epsilon}$, $\binom{w_i}{\epsilon}$, $\binom{\epsilon}{u_i}$, $\binom{\epsilon}{v_i}$, and $\binom{\epsilon}{w_i}$ where $0 < i \leq |c_3|$. The E_3- or E_4-edges leaving $\binom{u_i}{\epsilon}$, $\binom{v_i}{\epsilon}$, $\binom{w_i}{\epsilon}$, $\binom{\epsilon}{u_i}$, $\binom{\epsilon}{v_i}$, and $\binom{\epsilon}{w_i}$ are easily described.

We need to check that none of the generated vertices is of the form $\binom{c_1}{\epsilon}$, $\binom{c_2}{\epsilon}$, or $\binom{c_3}{\epsilon}$ (otherwise there would be a path from open to close). It is obvious that $\binom{c_1}{\epsilon}$ is not the end vertex of any E_3-edge and $\binom{c_3}{\epsilon}$ is not the end vertex of any E_4-edge. If $\binom{c_1}{\epsilon}$ is the end vertex of an E_4-edge of the form $\left(\binom{\epsilon}{u_i}, \binom{c_1}{\epsilon}\right)$ (resp., $\left(\binom{\epsilon}{v_i}, \binom{c_1}{\epsilon}\right)$, $\left(\binom{\epsilon}{w_i}, \binom{c_1}{\epsilon}\right)$), then $c_2 = u_i c_1$ or $c_3 = u_i c_1$ (resp., $c_2 = v_i c_1$ or $c_3 = v_i c_1$, $c_2 = w_i c_1$ or $c_3 = w_i c_1$). Similarly, if $\binom{c_2}{\epsilon}$ is the end vertex of an E_4-edge of the form $\left(\binom{\epsilon}{u_i}, \binom{c_2}{\epsilon}\right)$, respectively, $\left(\binom{\epsilon}{v_i}, \binom{c_2}{\epsilon}\right)$, $\left(\binom{\epsilon}{w_i}, \binom{c_2}{\epsilon}\right)$, then $c_3 = u_i c_2$ (resp., $c_3 = v_i c_2, c_3 = w_i c_2$). If $\binom{c_2}{\epsilon}$ is the end vertex of an E_3-edge of the form $\left(\binom{u_i}{\epsilon}, \binom{u_i c_1}{\epsilon}\right)$, respectively, $\left(\binom{v_i}{\epsilon}, \binom{v_i c_1}{\epsilon}\right)$, $\left(\binom{w_i}{\epsilon}, \binom{w_i c_1}{\epsilon}\right)$, then $c_2 = u_i c_1$ (resp., $c_2 = v_i c_1$, $c_2 = w_i c_1$). If $\binom{c_3}{\epsilon}$ is the end vertex of an E_3-edge of the form $\left(\binom{u_i}{\epsilon}, \binom{u_i c_1}{\epsilon}\right)$ (resp., $\left(\binom{v_i}{\epsilon}, \binom{v_i c_1}{\epsilon}\right)$, $\left(\binom{w_i}{\epsilon}, \binom{w_i c_1}{\epsilon}\right)$, $\left(\binom{u_i}{\epsilon}, \binom{u_i c_2}{\epsilon}\right)$, $\left(\binom{v_i}{\epsilon}, \binom{v_i c_2}{\epsilon}\right)$, $\left(\binom{w_i}{\epsilon}, \binom{w_i c_2}{\epsilon}\right)$), then $c_3 = u_i c_1$ (resp., $c_3 = v_i c_1$, $c_3 = w_i c_1$, $c_3 = u_i c_2$, $c_3 = v_i c_2$, $c_3 = w_i c_2$). We need to consider the cases $c_2 = v_i c_1$, $c_3 = w_i c_1$, and $c_3 = w_i c_2$.

For the rest of the discussion, we can assume that u or v is a prefix of c_1 (otherwise, it is not difficult to see that there is no path in $G(C)$ from *open* to *close*). Since u is not a prefix of v and $|u| < |v|$, u and v cannot be both prefixes of c_1. In either case, $|u| < |c_1|$. To simplify the notation, put $x_{|u|} = x$, and whenever $|v| < |c_1|$, put $x_{|v|} = y$.

First, assume that v is a prefix of c_1. Here, $y \neq \epsilon$ and $vy \neq yv$ by Lemma 1.4. To simplify the notation, put $u_{|u|} = w$. in the case where $c_3 = w_i c_1$, since $i = |c_1^{k-1} v|$, we have $w_i = c_1^{k-1} v$ and $c_3 = c_1^{k-1} v c_1 = c_1^k v$ yielding $vy = yv$, a contradiction. In the case where $c_3 = w_i c_2$, we have $i = |v| - |u|$, $w_i = u_i$, $c_3 = u_i c_1^k u = c_1^k v$, and $c_1 = u_i x_i = vy$. Put $v = u_i t$ where $t \neq \epsilon$. The equality $u_i c_1^k u = c_1^k u_i t$ implies that $t = u$ and $u_i x_i = x_i u_i$. But then, $c_1 = u_i x_i = u_i u y = u y u_i$. We then conclude that u is a prefix of c_1, which is a contradiction. In the case where $c_2 = v_i c_1$, we have $i = |c_1^{k-1} u|$. If $y \notin \text{Prefix}(C) \setminus \{\epsilon\}$, we get Figure 2.1.

Otherwise, put $vy = yv'$ where $v' \neq v$ ($v' \notin \text{Prefix}(C) \setminus \{\epsilon\}$). If u is not a prefix of v', then add the edge $\left(\binom{v}{\epsilon}, \binom{\epsilon}{y}\right)$ to Figure 2.1. Otherwise, put $v' = us$ where $s \neq \epsilon$. We have $v_i = c_1^{k-1} w$, $c_2 = c_1^{k-1} w c_1 = c_1^k u$, and therefore $c_1 = wx = xu = vy$. We have $v = wt$ where $t \neq \epsilon$, and therefore $wty = tyu = yus$ with $u \neq w$. By Lemma 1.5(1), we get $s \neq t$. Since $|s| = |t|$, we conclude that $s \notin \text{Prefix}(C) \setminus \{\epsilon\}$. Add the edges $\left(\binom{v}{\epsilon}, \binom{\epsilon}{y}\right)$ and $\left(\binom{c_1^{k-1} v}{\epsilon}, \binom{\epsilon}{yu}\right)$ to Figure 2.1.

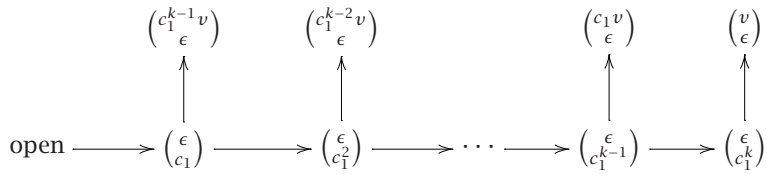

FIGURE 2.1

In the rest of the note, we assume that u is a prefix of c_1. Here, $x \neq \epsilon$ and $ux \neq xu$ by Lemma 1.4. In the case where $c_2 = v_i c_1$, we have $v_i = c_1^{k-1}u$ and $c_2 = c_1^{k-1}uc_1 = c_1^k u$. We conclude that $ux = xu$, a contradiction. In the cases where $c_3 = w_i c_1$ or $c_3 = w_i c_2$, if $x \notin \text{Prefix}(C) \setminus \{\epsilon\}$, we get Figure 2.2.

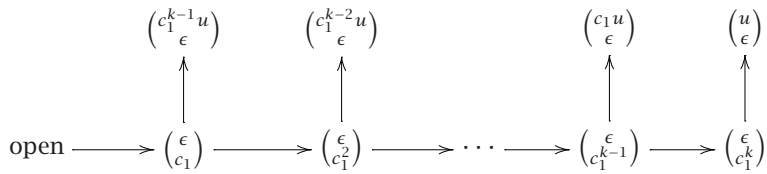

FIGURE 2.2

Otherwise, put $ux = xu'$ where $u' \neq u$ ($u' \notin \text{Prefix}(C) \setminus \{\epsilon\}$). If u' is not a prefix of v, then add the edge $\left(\binom{u}{\epsilon}, \binom{\epsilon}{x}\right)$ to Figure 2.2. Otherwise, put $v = u'c_1^q z$ where c_1 is not a prefix of z and $q \geq 0$. Note that if $z = \epsilon$, then $c_3 = c_1^k u' c_1^q = c_1^{k-1}uc_1^{q+1}$ and so $c_2 c_1^{q+1} = c_1 c_3$, a contradiction with the fact that C is SD. Note also that if $q \geq k$ and $z = u'$, then $c_1 c_3 = c_2 c_1^{q-k} c_2 c_1$, a contradiction with the fact that C is SD. It is not difficult to see that $c_1^i z \notin \text{Prefix}(C) \setminus \{\epsilon\}$ ($0 \leq i < k$) unless z is a prefix of c_1, and $c_1^k z \in \text{Prefix}(C) \setminus \{\epsilon\}$ only if z is a prefix of u or v. Whenever z is a prefix of c_1, put $x_{|z|} = x_1$ and whenever z is a prefix of u, put $c_2 = c_1^k z c_1^{p'} y'$ where $p' \geq 0$ and c_1 is not a prefix of y'. In the case where z is a prefix of v, put $c_3 = c_1^k z c_1^{q'} z'$ where $q' \geq 0$ and c_1 is not a prefix of z'.

The following points enable us to assume that $|z| < |u|$ and z is a prefix of u whenever z is a prefix of c_1 and $c_3 = w_i c_1$:

- if $|z| = |u|$, then $z = u = u'$, a contradiction;
- if $|z| > |u|$, then put $z = ut = t'u'$ fore some $t, t' \neq \epsilon$. Here, $c_1 = utx_1 = tx_1u' = t'u'x_1$ implies that $t = t'$. We also have $x = tx_1 = x_1 t$ and $c_1 = ux_1 t = x_1 tu' = tu'x_1$. These equalities cannot hold simultaneously by Lemma 1.5(1).

If $c_3 = w_i c_2$ and $|z| = |u|$, we have $z = u$ and z is not a prefix of v. In this case, if $q < k-1$, $u = u'$, a contradiction, and if $q \geq k-1$, C is not SD ($c_1 c_3 = c_2 c_1^{q-k+1} c_2$). The above and the following points enable us to assume that $|z| < |u|$, z is a prefix of u, and $q > k-1$ whenever z is a prefix of c_1 and $c_3 = w_i c_2$:

- if $|z| > |u|$, then put $z = ut = t'u$ for some $t, t' \neq \epsilon$. Here, $c_1 = utx_1 = tx_1u' = x_1u't'$ and $t = t'$. These equalities cannot hold simultaneously by Lemma 1.5(1);

- if $|z| < |u|$, then put $u = zt = t'z$ for some $t, t' \neq \epsilon$, and put $u' = z''t'$. In this case, if $q < k$, then $c_1^k u' c_1^q z = c_1^q z c_1^k u$ and $c_1 = ztx = txz'' = xz''t'$ with $t = t'$. These equalities cannot hold simultaneously by Lemma 1.5(1).

Trying to build a path from *open* to *close*, add $\left(\binom{u}{\epsilon}, \binom{\epsilon}{x}\right)$ and Figure 2.3 to Figure 2.2.

$$\binom{c_1^{k-1}u}{\epsilon} \longrightarrow \binom{c_1^k u'}{\epsilon} \longrightarrow \binom{c_1^k u' c_1}{\epsilon} \longrightarrow \cdots \longrightarrow \binom{c_1^k u' c_1^{q-1}}{\epsilon} \longrightarrow \binom{c_1^k u' c_1^q}{\epsilon}$$

FIGURE 2.3

The edge $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$ is added in case $q > k-1$, or $q = k-1$, u is a proper prefix of z, and z is not a prefix of v. The edge $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{c_1^k u' c_1^i c_2}{\epsilon}\right)$ is added in case $0 \leq i < q-k$. The edge $\left(\binom{c_1^k u' c_1^{q-k}}{\epsilon}, \binom{c_1^k u' c_1^{q-k} c_2}{\epsilon}\right)$ is added in case $q > k-1$, u is a proper prefix of z, and z is not a prefix of v. The edge $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{\epsilon}{c_1^{q+1}z}\right)$ is added in case $q < k-1$ and z is a prefix of u, or $q = k-1$ and z is a prefix of u or v. The edge $\left(\binom{c_1^k u' c_1^{q-k}}{\epsilon}, \binom{\epsilon}{c_1^k z}\right)$ is added in case $q > k-1$ and z is a prefix of u or v. If $i \geq 0$ and $q-k < i \leq q$, the edge $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{\epsilon}{c_1^{q-i} z}\right)$ is added in case z is a prefix of u. Note that no E_3- nor E_4-edge leaves $\binom{\epsilon}{x}$, and no E_3- nor E_4-edge leaves the vertices in Figure 2.3 other than the edges discussed above.

In the rest of the discussion, whenever u is a proper prefix of z, we put $z = ur$ where $r \neq \epsilon$ and x is not a prefix of r, and whenever u' is a prefix of z, we put $z = u'r'$. It is not difficult to see that $r \notin \mathrm{Prefix}(C) \setminus \{\epsilon\}$ (otherwise, $ur = z$ is a prefix of $ux = c_1$ and therefore z is a prefix of u, a contradiction). Note the following points:

- No E_3-edge leaves $\binom{c_1^{k-1}uc_2}{\epsilon}$ unless $q = k$ and u' is a prefix of z. No E_4-edge leaves $\binom{c_1^{k-1}uc_2}{\epsilon}$ other than $\left(\binom{c_1^{k-1}uc_2}{\epsilon}, \binom{\epsilon}{xz}\right)$ in case $q = k$ and $xz \in \mathrm{Prefix}(C) \setminus \{\epsilon\}$.
- No E_3- nor E_4-edges leave $\binom{c_1^k u' c_1^i c_2}{\epsilon}$ $(0 \leq i < q-k-1)$.
- For $q > k$ and $i = q-k-1$, no E_3-edge leaves $\binom{c_1^k u' c_1^i c_2}{\epsilon}$ unless u' is a prefix of z, and no E_4-edge leaves $\binom{c_1^k u' c_1^i c_2}{\epsilon}$ other than $\left(\binom{c_1^k u' c_1^i c_2}{\epsilon}, \binom{\epsilon}{xz}\right)$ in case $xz \in \mathrm{Prefix}(C) \setminus \{\epsilon\}$.
- No E_3- nor E_4-edges leave $\binom{c_1^k u' c_1^{q-k} c_2}{\epsilon}$.

We now state and prove our main results.

THEOREM 2.1. *Let C, u, v, and z be defined as described above. If z is a prefix of both u and v, then C is* UD.

PROOF. Here $|z| < |u|$. Put $u = zt$ and $u' = zt'$ where $t, t' \neq \epsilon$ and $t \neq t'$. First, assume that $c_3 = w_i c_1$. Here $u' = t'z$ and $c_1 = ztx = xzt' = xt'z = zxt' = txz$. By Lemma 1.5(2), the equalities $ztx = txz = xzt'$ cannot hold simultaneously. Now, assume that $c_3 = w_i c_2$. Here $u = t'z$ and $c_1 = ztx = xzt' = t'zx = zxt'$. Since $t'zx = zxt' = ztx$, by Lemma 1.6, $t = (s's)^\ell$, $t' = (ss')^\ell$, $x = (s's)^m s'$, $z = (ss')^n s$ for some $\ell > 0, m \geq 0, n \geq 0$, and words s, s' on A. The equality $ztx = xzt'$ implies that $ss' = s's$, and therefore $t = t'$, a contradiction. □

THEOREM 2.2. *Let C, u, v, and z be defined as described above. If z is a prefix of u but not a prefix of v, then C is* UD.

PROOF. Here, $|z| < |u|$. Put $u = zt$ and $u' = z''t'$, where $z \neq z''$ and $t, t' \neq \epsilon$. We have $xz \notin \mathrm{Prefix}(C) \setminus \{\epsilon\}$ since $c_1 = xz''t$ and $z \neq z''$.

First, assume that $c_3 = w_i c_1$. Here $u' = t'z$ and $c_1 = ztx = zxt' = txz = xz''t' = xt'z$. If $t = t'$, then $zxt = xtz = xz''t$. By Lemma 1.6, $z'' = (s's)^\ell$, $z = (ss')^\ell$, $t = (s's)^m s'$, $x = (ss')^n s$ for some $\ell > 0$, $m \geq 0$, $n \geq 0$, and words s, s' on A. The equality $txz = xtz$ implies that $ss' = s's$, and therefore $z = z''$, a contradiction. Otherwise, since $|t| < |u|$ and $c_1 = txz$, we get $u = tz'''$ for some z'''. We have $zx = xz''$ and $z'''x = xz$. Certainly $z''' \neq z$, and by Lemma 1.7, $z''' \neq z''$. By Lemma 1.6, since $ztx = txz = tz'''x$, put $z''' = (s's)^\ell$, $z = (ss')^\ell$, $x = (s's)^m s'$, and $t = (ss')^n s$ for some $\ell > 0$, $m \geq 0$, $n \geq 0$, and words s, s' on A. Whenever $m > 0$, the equality $ztx = xz''t'$ implies that $ss' = s's$ and $z = z'''$, a contradiction. Otherwise, x is a prefix of z''' and we put $z''' = xy$ and $z = yx$, where $y = (ss')^{\ell-1}s$. Here y cannot be a prefix of z'' or the equality $xz''t' = zxt'$ implies that $z = z'''$. Trying to build a path from *open* to *close*, the edge $\left(\binom{u}{\epsilon}, \binom{\epsilon}{x}\right)$ together with Figure 2.3 are added to Figure 2.2. We now discuss the other edges added.

First, assume that $q \leq k$. Whenever $q < k$, add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{\epsilon}{c_1^{q+1}z}\right)$ to the graph and whenever $q = k$, add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$. In either case, we then also add $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{\epsilon}{c_1^{q-i}z}\right)$ $(0 \leq i \leq q)$, $\left(\binom{\epsilon}{c_1^i z}, \binom{\epsilon}{c_1^{i+1}z}\right)$ $(0 \leq i < k)$, $\left(\binom{\epsilon}{z}, \binom{tx}{\epsilon}\right)$, $\left(\binom{tx}{\epsilon}, \binom{\epsilon}{z}\right)$, $\left(\binom{\epsilon}{c_1^i z}, \binom{c_1^{k-i}t}{\epsilon}\right)$ $(0 \leq i \leq k)$, $\left(\binom{c_1^i tx}{\epsilon}, \binom{c_1^{i+1}tx}{\epsilon}\right)$ $(0 \leq i < k)$, and $\left(\binom{c_1^i tx}{\epsilon}, \binom{\epsilon}{c_1^{k-i}y}\right)$ $(0 \leq i \leq k)$.

Second, assume that $q > k$. Add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$, $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{c_1^k u' c_1^i c_2}{\epsilon}\right)$ $(0 \leq i < q-k)$, and $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{\epsilon}{c_1^{q-i}z}\right)$ $(q-k \leq i \leq q)$. Also add $\left(\binom{\epsilon}{c_1^i z}, \binom{\epsilon}{c_1^{i+1}z}\right)$ $(0 \leq i < k)$, $\left(\binom{\epsilon}{z}, \binom{tx}{\epsilon}\right)$, $\left(\binom{tx}{\epsilon}, \binom{\epsilon}{z}\right)$, $\left(\binom{\epsilon}{c_1^i z}, \binom{c_1^{k-i}t}{\epsilon}\right)$ $(0 \leq i \leq k)$, $\left(\binom{c_1^i tx}{\epsilon}, \binom{c_1^{i+1}tx}{\epsilon}\right)$ $(0 \leq i < k)$, and $\left(\binom{c_1^i tx}{\epsilon}, \binom{\epsilon}{c_1^{k-i}y}\right)$ $(0 \leq i \leq k)$.

Now, assume that $c_3 = w_i c_2$. Here $u = t'z$ and $c_1 = ztx = t'zx = zxt' = xz''t' = t'xz''$. If $t = t'$, then $ztx = txz'' = xz''t$ and these equalities cannot hold simultaneously by Lemma 1.5(1). Otherwise, $t \neq t'$ and $t \notin \mathrm{Prefix}(C) \setminus \{\epsilon\}$. trying to build a path from *open* to *close*, the edge $\left(\binom{u}{\epsilon}, \binom{\epsilon}{x}\right)$ together with Figure 2.3 are added to Figure 2.2. We now discuss the other edges added.

First, assume that $q = k$. Add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$ and $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{\epsilon}{c_1^{q-i}z}\right)$ $(0 \leq i \leq q)$.

Second, assume that $q > k$. Add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$, $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{c_1^k u' c_1^i c_2}{\epsilon}\right)$ $(0 \leq i < q-k)$, and $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{\epsilon}{c_1^{q-i}z}\right)$ $(q-k \leq i \leq q)$. □

THEOREM 2.3. *Let C, u, v, and z be defined as described above. If z is a prefix of v but not a prefix of u, then C is* UD.

PROOF. Here $c_3 = c_1^k u' c_1^q z = c_1^k z c_1^{q'} z'$ and $|c_1^{q'} z'| = |c_1^q u|$. The latter and the fact that $|u| < |c_1|$ imply that $q \geq q'$. Note that u is not a prefix of z, u' is a suffix of z' in case $c_3 = w_i c_1$, and u is a suffix of z' in case $c_3 = w_i c_2$. Trying to build a path from *open* to *close*, the edge $\left(\binom{u}{\epsilon}, \binom{\epsilon}{x}\right)$ together with Figure 2.3 are added to Figure 2.2. We now discuss the other edges added (if any).

First, assume that $q < k-1$. In this case, no more edges are added.

Second, assume that $q = k-1$. We have $|c_1^{q'}z'| = |c_1^{k-1}u|$ and add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{\epsilon}{c_1^k z}\right)$ along with

$$\binom{\epsilon}{c_1^k z} \longrightarrow \binom{\epsilon}{c_1^k z c_1} \longrightarrow \cdots \longrightarrow \binom{\epsilon}{c_1^k z c_1^{q'-1}} \longrightarrow \binom{\epsilon}{c_1^k z c_1^{q'}}$$

FIGURE 2.4

In case $c_3 = w_i c_2$, we have $c_1^{q'}z' = c_1^{k-1}u$ ($q' = k-1$ and $z' = u$) and also add $\left(\binom{\epsilon}{c_1^k z c_1^{q'-i}}, \binom{c_1^i u}{\epsilon}\right)$ $(0 \le i < k)$.

Third, assume that $q = k$ and $c_3 = w_i c_2$. We have $c_1^{q'}z' = c_1^k u$ ($q' = k$ and $z' = u$) and $|z| \neq |u|$ here. If $|z| < |u|$, then $u' = zt$ and $u = tz$ for some $t \neq \epsilon$. The latter together with the equality $c_1^k u' c_1^k z = c_1^k z c_1^k u$ yield that $c_1 = tzx = xzt = zxt$ and therefore z is a prefix of u, a contradiction. If $|z| > |u|$, put $z = u'c_1^p t' = tu$ where $t' \neq \epsilon$ and c_1 is not a prefix of t nor t'. To simplify the notation, whenever $|t'| < |c_1|$, put $x_{|t'|} = x'$. From the equality $c_1^k u' c_1^k z = c_1^k z c_1^k u$, it follows that $c_1^k u' c_1^p t' = c_1^p t' c_1^k u$.

• If $p > k$, we contradict the fact that c_1 is not a prefix of z.

• If $p < k$, then $|t'| < |c_1|$ or we contradict the fact that c_1 is not a prefix of t' (t' is a prefix of c_1 here). If $|t'| < |u|$, put $u = t'r_1 = r_1't'$ and $u' = t''r_1'$. We then have $c_1 = t'r_1x = r_1't'x = xt''r_1' = r_1xt''$ and $r_1 = r_1'$. Then $t' \neq t''$ and we have $t'r_1x = r_1xt'' = xt''r_1$, which cannot hold simultaneously by Lemma 1.5(1). If $|t'| = |u|$, then $t' = u$ and, when $k-p = 1$, C is not SD ($c_1c_3 = c_2^3$). When $k-p > 1$, we get $u = u'$, a contradiction. So we have $|t'| > |u|$ and put $t' = ur_1 = r_1'u$. Then $c_1 = ur_1x' = r_1'ux' = r_1x'u' = x'u'r_1'$ and $r_1 = r_1'$. Consequently, $ur_1x' = r_1x'u' = x'u'r_1$, which cannot hold simultaneously by Lemma 1.5(1).

• If $p = k$ then from the equality $c_1^k u' c_1^p t' = c_1^p t' c_1^k u$ it follows that $u' c_1^k t' = t' c_1^k u$. Here, clearly $|t'| \neq |u|$. If $|t'| < |u|$, put $u' = t'r_1$ and $u = r_1t' = t''r_1'$ where $|t'| = |t''|$. This then yields that $c_1 = r_1t'x = t''r_1'x = xt'r_1 = t'xr_1$ and $t' = t''$. So we have $r_1 \neq r_1'$ and $r_1t'x = t'xr_1 = t'r_1'x$. Using Lemma 1.6 and the fact that $t'xr_1 = xt'r_1$, we conclude that $u = u'$, a contradiction. Thus $|t'| > |u|$ and put $t' = u'c_1^{p_1}t_1' = t_1u$ for some t_1 and t_1' where $t_1' \neq \epsilon$ and c_1 is not a prefix of t_1 nor t_1'. From the equality $u'c_1^k t' = t'c_1^k u$, it follows that $c_1^k u' c_1^{p_1} t_1' = c_1^{p_1} t_1' c_1^k u$ and we continue as above, comparing p_1 to k. Since c_3 is finite, we must reach $t_i' = u'c_1^{p_i+1}t_{i+1}' = t_{i+1}u$ where $t_{i+1}' \neq \epsilon$ and c_1 is not a prefix of t_{i+1} nor t_{i+1}', and consequently we get the equality $c_1^k u' c_1^{p_i+1} t_{i+1}' = c_1^{p_i+1} t_{i+1}' c_1^k u$. Then we have $p_{i+1} > k$, $p_{i+1} < k$, or $p_{i+1} = k$, and $|t_{i+1}'| \le |u|$. In all cases we reach the same types of contradictions as previously stated.

Fourth, assume that $q > k$ and $c_3 = w_i c_2$. Since $|c_1^q u| = c_1^{q'}z'$, it follows that c_2 is a suffix of $c_1^{q'}z'$ and we put $c_3 = c_1^k u' c_1^q z = c_1^k z c_1^{q'} z' = c_1^k z s c_1^k u$ where $|s| = |c_1^{q-k}|$. Clearly, $|z| \neq |u|$ here. If $|z| < |u|$, put $u' = zt$ and $u = tz$ for some $t \neq \epsilon$. The latter together with the equality $c_1^k u' c_1^q z = c_1^k z s c_1^k u$ yield that $c_1 = tzx = xzt = zxt$, which contradicts the fact that z is not a prefix of u. If $|z| > |u|$,

put $z = u'c_1^p t' = tu$ where $t' \neq \epsilon$ and c_1 is not a prefix of t nor t'. To simplify the notation, whenever $|t'| < |c_1|$, put $x_{|t'|} = x'$. From the equality $c_1^k u' c_1^q z = c_1^k z s c_1^k u$, it follows that $c_1^q u' c_1^p t' = c_1^p t' s c_1^k u$. Add $\left(\binom{c_1^{k-1}u}{\epsilon}, \binom{c_1^{k-1}uc_2}{\epsilon}\right)$, $\left(\binom{c_1^k u' c_1^{q-k}}{\epsilon}, \binom{\epsilon}{c_1^k z}\right)$, Figure 2.4, and $\left(\binom{c_1^k u' c_1^i}{\epsilon}, \binom{c_1^k u' c_1^i c_2}{\epsilon}\right)$ $(0 \leq i < q-k)$ along with $\left(\binom{c_1^k u' c_1^{q-k-1} c_2}{\epsilon}, \binom{c_1^k u' c_1^q u'}{\epsilon}\right)$ and Figure 2.5:

$$\binom{c_1^k u' c_1^q u'}{\epsilon} \longrightarrow \binom{c_1^k u' c_1^q u' c_1}{\epsilon} \longrightarrow \cdots \longrightarrow \binom{c_1^k u' c_1^q u' c_1^{p-1}}{\epsilon} \longrightarrow \binom{c_1^k u' c_1^q u' c_1^p}{\epsilon}$$

FIGURE 2.5

• If $p > q$, then we contradict the fact that c_1 is not a prefix of z.

• If $p < q$, then $|t'| < |c_1|$ or we contradict the fact that c_1 is not a prefix of t' (t' is a prefix of c_1 here). If $|t'| > |u|$, put $t' = ur_1 = r_1' u$. Then $c_1 = ur_1 x' = r_1' ux' = r_1 x' u' = x' u' r_1'$ and $r_1 = r_1'$. So $ur_1 x' = r_1 x' u' = x' u' r_1$, which cannot hold simultaneously by Lemma 1.5(1). If $|t'| = |u|$, then $t' = u$ and it follows from the equality $c_1^q u' c_1^p u = c_1^p usc_1^k u$ that $c_1^{q-p} u' c_1^p = usc_1^k$. Here if $k - p > 1$, we get $u = u'$, a contradiction. If $k - p \leq 1$ then C is not SD ($c_1 c_3 = c_2 c_1^{q-k} c_2 c_1^{p-k+1} c_2$). If $|t'| < |u|$, put $u = t' r_1 = r_1' t'$ and $u' = t'' r_1'$. We then have $c_1 = t' r_1 x = r_1' t' x = xt'' r_1' = t' x r_1' = r_1' x t''$. In case $k > p$, from the equality $c_1^q u' c_1^p t' = c_1^p t' s c_1^k u$, it follows that $c_1^{q-p} t'' = t' s c_1^{k-p}$. Consequently, we also have $c_1 = r_1 x t''$ and $r_1 = r_1'$. Then $t' \neq t''$ and $t' r_1 x = r_1 x t'' = xt'' r_1$, which cannot hold simultaneously by Lemma 1.5(1). So we have $p \geq k$, $r_1 \neq r_1'$, and $xz = c_1^{p+1} t' \notin \text{Prefix}(C) \setminus \{\epsilon\}$. Note that $q' = 0$ here or we have $r_1 = r_1'$. Also, $xt' \notin \text{Prefix}(C) \setminus \{\epsilon\}$ or we get $t' = t''$, which leads to a contradiction by Lemma 1.6. Add

$$\begin{gathered} \left(\binom{c_1^k u' c_1^q u' c_1^{p-i}}{\epsilon}, \binom{\epsilon}{c_1^i t'}\right) \quad (0 \leq i \leq k) \text{ along with} \\ \left(\binom{c_1^k u' c_1^q u' c_1^i}{\epsilon}, \binom{c_1^k u' c_1^q u' c_1^i c_2}{\epsilon}\right) \quad (0 \leq i < p-k) \end{gathered} \tag{2.1}$$

and $\left(\binom{c_1^k u' c_1^{q-k-1} c_2}{\epsilon}, \binom{c_1^k u' c_1^q u' c_1^{k-1} u}{\epsilon}\right)$.

• If $p = q$, from the equality $c_1^q u' c_1^p t' = c_1^p t' s c_1^k u$, it follows that $u' c_1^q t' = t' s c_1^k u$. Clearly, $|t'| \neq |u|$, and if $|t'| < |u|$,then put $u' = t' r_1$ and $u = r_1 t' = t'' r_1'$ where $|t'| = |t''|$. We then have $r_1 t' x = t'' r_1' x = xt' r_1 = t' x r_1$ and $t' = t''$. Consequently, $r_1 \neq r_1'$ and we get $r_1 t' x = t' x r_1 = t' r_1' x$. Using Lemma 1.6 and the fact that $xt' r_1 = t' x r_1$, we reach a contradiction with the fact that $u \neq u'$. So we have $|t'| > |u|$ and put $t' = u' c_1^{p_1} t_1' = t_1 u$, where $t_1' \neq \epsilon$ and c_1 is not a prefix of t_1 nor t_1'. From the equality $u' c_1^q t' = t' s c_1^k u$, it follows that $c_1^q u' c_1^{p_1} t_1' = c_1^{p_1} t_1' s c_1^k u$ and we continue as above, comparing p_1 with q. Since c_3 is finite, we must reach the equality $t_i' = u' c_1^{p_{i+1}} t_{i+1}' = t_{i+1}' u$ where $t_{i+1}' \neq \epsilon$ and c_1 is not a prefix of t_{i+1} nor t_{i+1}', and consequently we get the equality $c_1^q u' c_1^{p_{i+1}} t_{i+1}' = c_1^{p_{i+1}} t_{i+1}' s c_1^k u$. Then one of the following must occur:

(1) $p_{i+1} > q$, thus yielding the same type of contradiction as above;

(2) $p_{i+1} < q$ and $|t_{i+1}'| \geq |u|$, thus yielding the same types of contradiction as above;

(3) $p_{i+1} < q$, $|t'_{i+1}| < |u|$, and $p_{i+1} < k$, thus yielding the same type of contradiction as above;

(4) $p_{i+1} < q$, $|t'_{i+1}| < |u|$, and $p_{i+1} \geq k$, in which case we add

$$\left(\begin{pmatrix} c_1^k u' (c_1^q u')^{i+2} c_1^{p_i+1-j} \\ \epsilon \end{pmatrix}, \begin{pmatrix} \epsilon \\ c_1^j t'_{i+1} \end{pmatrix}\right) \quad (0 \leq j \leq k),$$

$$\left(\begin{pmatrix} c_1^k u' (c_1^q u')^{i+2} c_1^j \\ \epsilon \end{pmatrix}, \begin{pmatrix} c_1^k u' (c_1^q u')^{i+2} c_1^j c_2 \\ \epsilon \end{pmatrix}\right) \quad (0 \leq j < p_{i+1} - k),$$

$$\left(\begin{pmatrix} c_1^k u' (c_1^q u')^{i+1} c_1^{q-k-1} c_2 \\ \epsilon \end{pmatrix}\right), \left(\begin{pmatrix} c_1^k u' (c_1^q u')^{i+2} c_1^{k-1} u \\ \epsilon \end{pmatrix}\right); \tag{2.2}$$

(5) $p_{i+1} = q$ and $|t'_{i+1}| \leq |u|$, thus yielding the same types of contradiction as above.

Last, assume that $q \geq k$ and $c_3 = w_i c_1$. In this case, add $\left(\begin{pmatrix} c_1^{k-1} u \\ \epsilon \end{pmatrix}, \begin{pmatrix} c_1^{k-1} u c_2 \\ \epsilon \end{pmatrix}\right)$, $\left(\begin{pmatrix} c_1^k u' c_1^{q-k} \\ \epsilon \end{pmatrix}, \begin{pmatrix} \epsilon \\ c_1^k z \end{pmatrix}\right)$, and Figure 2.4. Also, add the edges $\left(\begin{pmatrix} c_1^k u' c_1^i \\ \epsilon \end{pmatrix}, \begin{pmatrix} c_1^k u' c_1^i c_2 \\ \epsilon \end{pmatrix}\right)$ $(0 \leq i < q-k)$.

If $|z| < |u|$ put $u' = zt' = t'z$ and $u = z''t''$ where $|t'| = |t''|$, $z \neq z''$, and $t', t'' \neq \epsilon$. Then we have $c_1 = xzt' = xt'z = z''t''x = t''xz$. Here, if $t' = t''$ we have the equalities $z''t'x = t'xz = xzt'$, which cannot hold by Lemma 1.5(1). So $t' \neq t''$ and we add $\left(\begin{pmatrix} c_1^{k-1} u c_2 \\ \epsilon \end{pmatrix}, \begin{pmatrix} \epsilon \\ xz \end{pmatrix}\right)$ whenever $q = k$ and $\left(\begin{pmatrix} c_1^k u' c_1^{q-k-1} c_2 \\ \epsilon \end{pmatrix}, \begin{pmatrix} \epsilon \\ xz \end{pmatrix}\right)$ whenever $q > k$. Note that here $q' = 0$ nor we contradict the fact that $t' \neq t''$.

If $|z| = |u|$ then $z = u'$ and C is not SD ($c_1 c_3 = c_2 c_1^{q-k} c_2 c_1$).

If $|z| > |u|$ put $z = u' c_1^p t' = tu'$ where c_1 is not a prefix of t or t', and $t' \neq \epsilon$ (otherwise, $c_1 c_3 = c_1 c_1^k u' c_1^q u' c_1^p = c_2 c_1^{q-k} c_2 c_1^{p+1}$). To simplify the notation, whenever $|t'| < |c_1|$, put $x_{|t'|} = x'$. Then we have $c_3 = c_1^k u' c_1^q z = c_1^k z c_1^{q'} z' = c_1^k z s c_1$, where $|s| = |c_1^{q-1} u|$ and it follows that $c_1^q u' c_1^p t' = c_1^p t' s c_1$. Here, we add $\left(\begin{pmatrix} c_1^{k-1} u c_2 \\ \epsilon \end{pmatrix}, \begin{pmatrix} c_1^k u' c_1^q u' \\ \epsilon \end{pmatrix}\right)$ if $q = k$, and $\left(\begin{pmatrix} c_1^k u' c_1^{q-k-1} c_2 \\ \epsilon \end{pmatrix}, \begin{pmatrix} c_1^k u' c_1^q u' \\ \epsilon \end{pmatrix}\right)$ if $q > k$. In either case, also add Figure 2.5.

- If $p > q$, then we contradict the fact that c_1 is not a prefix of z.
- If $p < q$, then $|t'| < |c_1|$ or we contradict the fact that c_1 is not a prefix of t' (t' is a prefix of $c_1$here). Clearly, $|t'| \neq |u|$ or we contradict the fact that $u \neq u'$. If $|t'| > |u|$, put $t' = ur_1 = r'_1 u'$ and $c_1 = ur_1 x' = r'_1 u' x' = r_1 x' u' = x' r'_1 u' = ux' r'_1$. Then $ux' r'_1 = x' r'_1 u' = r'_1 u' x'$, which cannot hold simultaneously by Lemma 1.5(1). If $|t'| < |u|$, put $u = t' r_1 = r''_1 t'''$ and $u' = r'_1 t' = t'' r'_1$ where $|t'| = |t'''|$. Then we have $c_1 = t' r_1 x = x r'_1 t' = t' x r'_1 = r_1 x t' = x t'' r'_1 = r''_1 t''' x$ and $r_1 = r''_1$. Here, if $t' = t''$, then $r_1 \neq r'_1$ and we have $t' r_1 x = r_1 x t' = x t' r'_1$, which cannot hold simultaneously by Lemma 1.5(2). So we assume that $t' \neq t''$; and if $r_1 = r'_1$, then the equalities $t' x r_1 = x r_1 t' = x t'' r_1$ together with Lemma 1.6 and the fact that $x r_1 t' = r_1 x t'$ yield a contradiction with the fact that $u \neq u'$. Also, if $t' = t'''$ then $x t' = t' x = x t''$ and $t' = t''$, a contradiction. So we have $t' \neq t''$, $t' \neq t'''$, $r_1 \neq r'_1$, and $xz = c_1^{p+1} t'$. Here, note that $q' \leq q - p - 1$ or we contradict the fact that $t' \neq t''$. Thus $|z'| = |c_1^{q-q'} u| > |u|$, r_1 is a prefix of z', and u' cannot be a prefix of z'.

In the case where $q = k$, we add

$$\left(\binom{c_1^{k-1}uc_2}{\epsilon},\binom{\epsilon}{xz}\right),\quad \left(\binom{c_1^k u' c_1^q u' c_1^i}{\epsilon},\binom{\epsilon}{c_1^{p-i}t'}\right)\quad (0 \le i \le p). \tag{2.3}$$

In the case where $q > k$ and $p < k$, add $\left(\binom{c_1^k u' c_1^{q-k-1} c_2}{\epsilon},\binom{\epsilon}{xz}\right)$ with (2.3). The edge

$$\left(\binom{\epsilon}{c_1^k z c_1^i},\binom{\epsilon}{c_1^k z c_1^i c_2}\right) \tag{2.4}$$

is added in case $0 \le i < q' - k$, or $i = q' - k$ and u is a prefix of z'.

In the case where $q > k$ and $p \ge k$, in addition to (2.1) and $\left(\binom{c_1^k u' c_1^{q-k-1} c_2}{\epsilon},\right.$ $\left.\binom{c_1^k u' c_1^q u' c_1^{k-1} u}{\epsilon}\right)$, add $\left(\binom{c_1^k u' c_1^q u' c_1^{p-i}}{\epsilon},\binom{\epsilon}{c_1^i t'}\right)$ $(0 \le i \le k)$. Also, (2.4) is added in case $0 \le i < q' - k$, or $i = q' - k$ and u is a prefix of z'.

In all cases, also add $\left(\binom{\epsilon}{c_1^i t'},\binom{\epsilon}{c_1^{i+1} t'}\right)$ $(0 \le i < k)$ and $\left(\binom{\epsilon}{c_1^i t'},\binom{c_1^{k-i} r_1}{\epsilon}\right)$ $(0 \le i \le k)$ as well as $\left(\binom{\epsilon}{t'},\binom{r_1 x}{\epsilon}\right)$, $\left(\binom{r_1 x}{\epsilon},\binom{\epsilon}{t'}\right)$, and $\left(\binom{c_1^i r_1 x}{\epsilon},\binom{c_1^{i+1} r_1 x}{\epsilon}\right)$ $(0 \le i < k-1)$. In the case where $|t'''| > |x|$, put $t''' = xs$ and $t' = sx = xs'$. It follows from the equality $t' x r_1' = x t'' r_1'$ that $t'' = s'x$ and consequently, $s \ne s'$. Add $\left(\binom{c_1^{k-1} r_1 x}{\epsilon},\binom{c_1^k r_1 x}{\epsilon}\right)$ along with $\left(\binom{c_1^i r_1 x}{\epsilon},\binom{\epsilon}{c_1^{k-i} s}\right)$ $(0 \le i \le k)$.

• If $p = q$, then it follows from $c_1^q u' c_1^p t' = c_1^p t' s c_1$ that $u' c_1^q t' = t' s c_1$. Here, if $|t'| = |u|$, then $t' = u'$ and C is not SD ($c_1 c_3 = c_2 c_1^{q-k} c_2 c_1^{q-k} c_2 c_1$). If $|t'| < |u|$ then put $u' = t' r_1 = r_1 t'$ and $u = t'' r_1'$, where $|t'| = |t''|$. So we have $c_1 = x t' r_1 = x r_1 t' = t'' r_1' x = r_1' x t'$. If $t' = t''$, then $r_1 \ne r_1'$ and we get $t' r_1' x = r_1' x t' = x t' r_1$, which cannot hold simultaneously by Lemma 1.5(2). So we assume that $t' \ne t''$, and if $r_1 = r_1'$ then $t'' r_1 x = r_1 x t' = x t' r_1$, which cannot hold by Lemma 1.5(1). Thus $t' \ne t''$, $r_1 \ne r_1'$, and $xz = x u' c_1^q t' = c_1^{q+1} t' \notin \text{prefix}(C) \setminus \{\epsilon\}$. Note that here $q' = 0$ or we contradict the fact that $r_1 \ne r_1'$. When $q = k$, we add $\left(\binom{c_1^{k-1} u c_2}{\epsilon},\binom{c_1^k u' c_1^k u' c_1^{k-1} u}{\epsilon}\right)$ as well as $\left(\binom{c_1^k u' c_1^k u' c_1^{k-1} u}{\epsilon},\binom{\epsilon}{xt'}\right)$ and $\left(\binom{c_1^k u' c_1^k u'}{\epsilon},\binom{\epsilon}{c_1^k t'}\right)$. When $q > k$, add (2.1) and $\left(\binom{c_1^k u' c_1^{q-k-1} c_2}{\epsilon},\binom{c_1^k u' c_1^q u' c_1^{k-1} u}{\epsilon}\right)$ along with $\left(\binom{c_1^k u' c_1^q u' c_1^{q-k-1} c_2}{\epsilon},\binom{\epsilon}{xt'}\right)$ and $\left(\binom{c_1^k u' c_1^q u' c_1^{q-k}}{\epsilon},\binom{\epsilon}{c_1^k t'}\right)$. If $|t'| > |u|$, put $t' = u' c_1^{p_1} t_1' = t_1 u'$, where c_1 is not a prefix of t_1 nor t_1', and $t_1' \ne \epsilon$ (otherwise, $c_1 c_3 = c_1 c_1^k u' (c_1^q u')^2 c_1^{p_1} = (c_2 c_1^{q-k})^2 c_2 c_1^{p_1+1}$). From the equality $u' c_1^q t' = t' s c_1$, it follows that $c_1^q u' c_1^{p_1} t_1' = c_1^{p_1} t_1' s c_1$ and we continue as above, comparing p_1 to q. Since c_3 is finite, we must reach $t_i' = u' c_1^{p_{i+1}} t_{i+1}' = t_{i+1} u'$ where c_1 is not a prefix of t_{i+1} or t_{i+1}', and $t_{i+1}' \ne \epsilon$ (otherwise, $c_1 c_3 = c_1 c_1^k u' (c_1^q u')^{i+2} c_1^{p_{i+1}} = (c_2 c_1^{q-k})^{i+2} c_2 c_1^{p_{i+1}+1}$), and consequently we get the equality $c_1^q u' c_1^{p_{i+1}} t_{i+1}' = c_1^{p_{i+1}} t_{i+1}' s c_1$. Then one of the following occurs:

(1) $p_{i+1} > q$, in which case we reach the same type of contradiction as above;
(2) $p_{i+1} < q$ and $|t_{i+1}'| \ge |u|$, in which case we reach the same type of contradictions as above;
(3) $p_{i+1} < q$ and $|t_{i+1}'| < |u|$, in which case we put $u = t_{i+t}' r_{i+2} = r_{i+2}'' t_{i+1}'''$ and $u' = r_{i+2}' t_{i+1}' = t_{i+1}'' r_{i+2}'$ where $|t_{i+1}'| = |t_{i+1}'''|$. As above, we have $t_{i+1}' \ne t_{i+1}''$, $t_{i+1}' \ne t_{i+1}'''$, $r_{i+2} \ne r_{i+2}'$, and $xz = c_1^{p_{i+1}+1} t_{i+1}' \notin \text{prefix}(C) \setminus \{\epsilon\}$. Also, $q' \le q - p_{i+1} - 1$, $|z'| = |c_1^{q-q'} u| > |u|$, r_{i+2} is a prefix of z', and u' cannot be a prefix of z'.

Add $\left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^j}{\epsilon}, \binom{\epsilon}{c_1^{p_{i+1}-j} t'_{i+1}}\right)$ $(0 \le j \le p_{i+1})$ whenever $q = k$, or $q > k$, and $p_{i+1} < k$. Whenever $q > k$ and $p_{i+1} \ge k$, add

$$\left(\binom{c_1^k u' (c_1^q u')^{i+1} c_1^{q-k-1} c_2}{\epsilon}, \binom{c_1^k u' (c_1^q u')^{i+2} c_1^{k-1} u}{\epsilon}\right),$$
$$\left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^j}{\epsilon}, \binom{c_1^k u' (c_1^q u')^{i+2} c_1^j c_2}{\epsilon}\right), \quad (0 \le j < p_{i+1} - k), \tag{2.5}$$

and $\left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^{p_{i+1}-j}}{\epsilon}, \binom{\epsilon}{c_1^j t'_{i+1}}\right)$ $(0 \le j \le k)$. In all cases, we also include $\left(\binom{\epsilon}{c_1^j t'_{i+1}}, \binom{\epsilon}{c_1^{j+1} t'_{i+1}}\right)$ $(0 \le j < k)$, $\left(\binom{\epsilon}{c_1^j t'_{i+1}}, \binom{c_1^{k-j} r_{i+2}}{\epsilon}\right)$ $(0 \le j \le k)$, $\left(\binom{\epsilon}{t'_{i+1}}, \binom{r_{i+2}x}{\epsilon}\right)$, $\left(\binom{r_{i+2}x}{\epsilon}, \binom{\epsilon}{t'_{i+1}}\right)$, and $\left(\binom{c_1^j r_{i+2}x}{\epsilon}, \binom{c_1^{j+1} r_{i+2}x}{\epsilon}\right)$ $(0 \le j < k-1)$, In the case where $q > k$, add $\left(\binom{\epsilon}{c_1^k z c_1^j}, \binom{\epsilon}{c_1^k z c_1^j c_2}\right)$ whenever $0 \le j < q' - k$, or $j = q' - k$ and u is a prefix of z'. In the case where $|t'''_{i+1}| > |x|$, put $t''''_{i+1} = x s_{i+1}$, $t'_{i+1} = s_{i+1}x = x s'_{i+1}$, and $t''_{i+1} = s'_{i+1}x$ where $s_{i+1} \ne s'_{i+1}$. Add $\left(\binom{c_1^{k-1} r_{i+2}x}{\epsilon}, \binom{c_1^k r_{i+2}x}{\epsilon}\right)$ along with $\left(\binom{c_1^j r_{i+2}x}{\epsilon}, \binom{\epsilon}{c_1^{k-j} s_{i+1}}\right)$ $(0 \le j \le k)$;

(4) $p_{i+1} = q$ and $|t'_{i+1}| = |u|$, in which case we reach the same type of contradiction as above;

(5) $p_{i+1} = q$ and $|t'_{i+1}| < |u|$, in which case, when $q = k$, we add

$$\left(\binom{(c_1^k u')^{i+1} c_1^{k-1} u c_2}{\epsilon}, \binom{(c_1^k u')^{i+3} c_1^{k-1} u}{\epsilon}\right), \quad \left(\binom{(c_1^k u')^{i+3}}{\epsilon}, \binom{\epsilon}{c_1^k t'_{i+1}}\right), \tag{2.6}$$

and $\left(\binom{(c_1^k u')^{i+3} c_1^{k-1} u}{\epsilon}, \binom{\epsilon}{x t'_{i+1}}\right)$. Whenever $q > k$, we add

$$\left(\binom{c_1^k u' (c_1^q u')^{i+1} c_1^{q-k-1} c_2}{\epsilon}, \binom{c_1^k u' (c_1^q u')^{i+2} c_1^{k-1} u}{\epsilon}\right),$$
$$\left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^j}{\epsilon}, \binom{c_1^k u' (c_1^q u')^{i+2} c_1^j c_2}{\epsilon}\right), \quad (0 \le j < q - k), \tag{2.7}$$
$$\left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^{q-k}}{\epsilon}, \binom{\epsilon}{c_1^k t'_{i+1}}\right), \quad \left(\binom{c_1^k u' (c_1^q u')^{i+2} c_1^{q-k-1} c_2}{\epsilon}, \binom{\epsilon}{x t'_{i+1}}\right).$$

□

Acknowledgments. This material is based upon work supported by the National Science Foundation (NSF) under Grant No. CCR-9700228. We thank the referees of a preliminary version of this note for their very valuable comments and suggestions.

References

[1] A. Apostolico and R. Giancarlo, *Pattern matching machine implementation of a fast test for unique decipherability*, Inform. Process. Lett. **18** (1984), no. 3, 155–158.

[2] F. Blanchet-Sadri, *On unique, multiset, and set decipherability of three-word codes*, IEEE Trans. Inform. Theory **47** (2001), no. 5, 1745–1757.
[3] F. Guzmán, *A complete list of small proper MSD and SD codes*, in preparation.
[4] _____, *Decipherability of codes*, J. Pure Appl. Algebra **141** (1999), no. 1, 13–35.
[5] T. Head and A. Weber, *Deciding multiset decipherability*, IEEE Trans. Inform. Theory **41** (1995), no. 1, 291–297.
[6] A. Hoffmann, *A test on unique decipherability*, MFCS 84, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1984, pp. 50–63.
[7] A. Lempel, *On multiset decipherable codes*, IEEE Trans. Inform. Theory **32** (1986), no. 5, 714–716.
[8] M. Lothaire, *Combinatorics on Words*, Encyclopedia of Mathematics and Its Applications, vol. 17, Addison-Wesley, 1983.
[9] M. Rodeh, *A fast test for unique decipherability based on suffix trees*, IEEE Trans. Inform. Theory **28** (1982), 648–651.

F. Blanchet-Sadri and T. Howell: Department of Mathematical Sciences, University of North Carolina, P.O. Box 26170, Greensboro, NC 27402-6170, USA