On Universal Cycles for Multisets

Glenn Hurlbert[*]
Tobias Johnson[†]
Joshua Zahl[‡]

January 17, 2007

Abstract

A *Universal Cycle* for t-multisets of $[n] = \{1, \ldots, n\}$ is a cyclic sequence of $\binom{n+t-1}{t}$ integers from $[n]$ with the property that each t-multiset of $[n]$ appears exactly once consecutively in the sequence. For such a sequence to exist it is necessary that n divides $\binom{n+t-1}{t}$, and it is reasonable to conjecture that this condition is sufficient for large enough n in terms of t. We prove the conjecture completely for $t \in \{2, 3\}$ and partially for $t \in \{4, 6\}$.

[*]Department of Mathematics and Statistics, Arizona State University, Tempe, Arizona 85287-1804, `hurlbert@asu.edu`

[†]Department of Mathematics, Yale University, New Haven, CT 06520, `tobias.l.johnson@gmail.com`. Research supported in part by NSF grant 0552730.

[‡]Department of Mathematics, Caltech, Pasadena, CA 91125, `jzahl@zahl.ca`. Research supported in part by NSF grant 0552730.

1 Introduction

Problem 109 in Section 7.2.1.3 of Donald Knuth's *The Art of Computer Programming* [4] lists the following sequence S.

$$111233522234413334552444511355551224$$

The noteworthy property of this sequence is discovered by listing every consecutive triple, including the two formed by wrapping the sequence cyclically. Here we have

$$111, 112, 123, \ldots, 224, 241, 411\ .$$

Not only are the triples distinct, but they are still distinct when considered as unordered multisets. In fact, each 3-multiset of $[5] = \{1, 2, 3, 4, 5\}$ appears exactly once. Such a sequence is called a *Universal Cycle for* 3*-multisets of* [5]. In this paper we will use the shortened term *Mcycle*, and in particular *t-Mcycle*, to refer to a universal cycle for t-multisets of $[n]$, just as the term *Ucycle* has come to refer to a Universal Cycle for t-subsets of $[n]$. Universal Cycles for a wide range of combinatorial structures were introduced in [1], and [2] contains the most up-to-date knowledge on Ucycles for subsets (not to mention the more recent and encouraging particular results of [3]). In this work we concern ourselves with t-Mcycles for $t \in \{3, 4, 6\}$. Note that any permutation of $[n]$ is a 1-Mcycle, and any Eulerian circuit of the complete graph on n vertices (with loops) is a 2-Mcycle. The main conjecture is the following.

Conjecture 1 *For t large enough in terms of n, Universal Cycles for t-multisets of $[n]$ exist if and only if n divides $\binom{n+t-1}{t}$.*

That the condition above is necessary follows from the fact that each symbol is in the same number of multisets and hence must appear equally often in the cycle. Our preceding comments indicate that the conjecture is true for $t \in \{1, 2\}$. In Section 2 we prove the following theorem.

Theorem 2 *Let $n_0(3) = 4$, $n_0(4) = 5$ and $n_0(6) = 11$. Then, for $t \in \{3, 4, 6\}$ and $n \geq n_0(t)$, Mcycles for t-multisets of $[n]$ exist whenever n is relatively prime to t.*

This theorem verifies the conjecture for $t = 3$, but leaves open the case $n \equiv 2 \mod 4$ for $t = 4$ and many cases for $t = 6$.

Because of the difficulty of extending these results for other values of t, it is worth considering other methods of construction. In Section 3 we describe an inductive technique for the case $t = 3$. This inductive technique has potential for generalization to higher values of t and it itself is an extension of a method developed by Anant Godbole and presented at the 2004 Banff conference on Generalizations of de Bruijn Cycles and Gray Codes. In Section 4 we outline a technique to convert 3-Ucycles into 3-Mcycles which also has potential for generalization.

2 Proof of Theorem 2

We use the techniques and terminology of [2] with one small exception of language (changing "d-set" to "form", below).

Let $S = \{s_1, \ldots, s_t\}$, $s_i \leq s_{i+1}$, be a t-multiset of $[n]$. Define its *form*, $F(s) = (f_1, \ldots, f_t)$ by $f_i = s_{i+1} - s_i$, where indices are computed modulo t and arithmetic is performed modulo n (where n is used in place of 0 as the modular representative). Two forms are *equivalent* if one is a cyclic permutation of the other. Two forms belong to the same *class* whenever one is any permutation of the other. We see that the collection of all classes is the collection of all unordered partitions of the integer n into t parts. Each class defines a partition of t according to the number of parts of the same size. Two classes belong to the same *pattern* if they define the same partition of t. We say that a pattern is *good* if some part has size 1 and *bad* otherwise. It may be useful to look at a few examples with $t = 5$ and $n = 30$.

In this case there are 7 patterns:

$$\langle 1,1,1,1,1\rangle, \langle 2,1,1,1\rangle, \langle 3,1,1\rangle, \langle 4,1\rangle, \langle 5\rangle, \langle 2,2,1\rangle, \text{ and } \langle 3,2\rangle \ .$$

Only $\langle 5\rangle$ and $\langle 3,2\rangle$ are bad. The pattern $\langle 3,2\rangle$ contains the 5 classes:

$$[0,0,0,15,15], [2,2,2,12,12], [4,4,4,9,9], [8,8,8,3,3], \text{ and } [10,10,10,0,0] \ .$$

(Note that the class $[6,6,6,6,6]$ is skipped from this sequence because it belongs to the pattern $\langle 5\rangle$.) The class $[10,10,10,0,0]$ contains the two forms

$$(10,10,10,0,0) \text{ and } (10,10,0,10,0) \ .$$

Finally, the form $(10,10,0,10,0)$ contains the 30 multisets

$$\{1,11,21,21,1\}, \{2,12,22,22,2\}, \ldots, \{30,10,20,20,30\} \ .$$

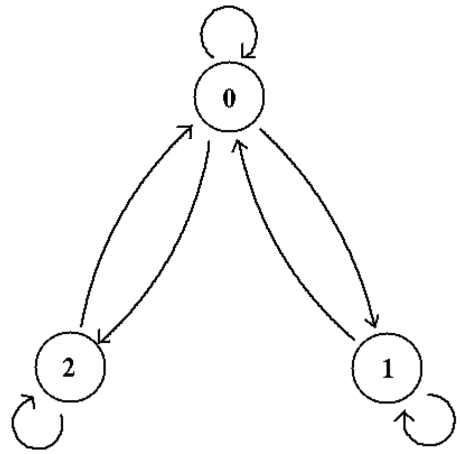

Figure 1: The transition graph $\mathcal{T}_{5,3}$

We will maintain the use of braces, parentheses, brackets, and angles in order to distinguish the various objects from one another.

We begin the proof by choosing a representative for each class. We distinguish one of the coordinates of the form $(f_1, \ldots, f_{t-1}, f_t)$ (because of the equivalence among its cyclic permutations we may assume it to be f_t) in the representation $(f_1, \ldots, f_{t-1}; f_t)$ so as to infer the ordering $\{i, i + f_1, \ldots, i + f_1 + \cdots + f_{t-1}\}$ of all its multisets. This singled out coordinate is therefore unused in the linear listing of each of these multisets.

Similarly, we may represent a class by $[f_1, \ldots, f_{t-1}; f_t]$, signifying that f_t is distinguished (unused) in each of its forms. It is important, then, that f_t be unique in order to avoid ambiguity—this is the reason for wanting good patterns. For example, with $t = 4$ and $n = 7$, we can choose $[1, 1, 0; 5]$ to represent [1,1,0,5]. This determines the representations $(1, 1, 0; 5), (1, 0, 1; 5)$ and $(0, 1, 1; 5)$ of its three forms, of which $(1, 0, 1; 5)$ denotes the (ordered) forms $\{1, 2, 2, 3\}$, $\{2, 3, 3, 4\}, \ldots$ and $\{7, 1, 1, 2\}$.

Based on these choices we define the *transition graph* $\mathcal{T}_{n,t}$ as follows. We define the *prefix*, resp. *suffix*, of the form representation $(f_1, \ldots, f_{t-1}; f_t)$ to be $((f_1, \ldots, f_{t-2}))$, resp. $((f_2, \ldots, f_{t-1}))$. Our use of double parentheses denotes that these are the vertices in the transition graph $\mathcal{T}_{n,t}$ whose directed edges are precisely the representations involved.

For example, Figure 1 shows the transition graph $\mathcal{T}_{5,3}$, which was used to construct the Mcycle for 3-multisets at the beginning of the article. The

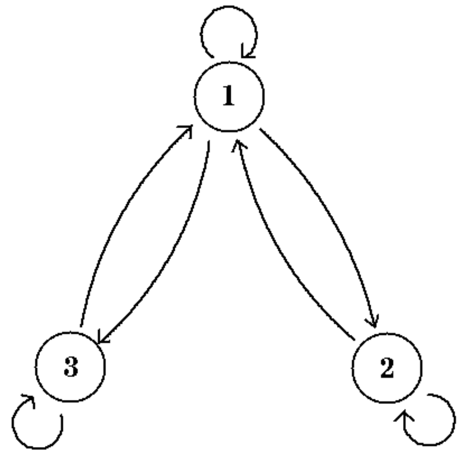

Figure 2: The transition graph $\mathcal{G}_{8,3}$

forms are represented by $(0,0;5),(1,1;3),\ (2,2;1),(0,1;4),\ (1,0;4),(0,2;3)$ and $(2,0;3)$. The form $(1,0;4)$ corresponds to the directed edge $((1)) \rightarrow ((0))$, and so on. The Eulerian circuit 0011022 corresponds to a listing of all forms and produces the differences in the first block, 0001224, along with the first digit, 1, of the next block. Since the sum $0+0+1+1+0+2+2 \equiv 1$ mod 5, each block shifts by 1, and since 1 is relatively prime to 5 each integer occurs as the starting digit of some block. Hence, each 3-multiset of [5] occurs exactly once. It turns out, however (see [2]), that having the sum relatively prime to n is an unnecessary component in the general construction process.

In [2] the analogous transition graph $\mathcal{G}_{n,t}$ is defined for t-subsets of $[n]$. For example, Figure 2 shows the transition graph $\mathcal{G}_{8,3}$, which can be used to construct a Ucycle for 3-subsets of [8]. The forms in this case are represented by $(1,1;6),(2,2;4),\ (3,3;2),(1,2;5),\ (2,1;5),(1,3;4)$ and $(3,1;4)$. The Eulerian circuit 1122133 generates a Ucycle as above.

In [2] we find the following fact.

Fact 3 *If $\mathcal{G}_{n,t}$ is Eulerian for some choice of representations of classes, then there exists a Ucycle for t-subsets of $[n]$.*

The same arguments that prove Fact 3 yield the analogous result for Mcycles, which we therefore state without proof.

Lemma 4 *If $\mathcal{T}_{n,t}$ is Eulerian for some choice of representations of classes, then there exists an Mcycle for t-multisets of $[n]$.*

The key is that the obvious isomorphism between $\mathcal{T}_{5,3}$ and $\mathcal{G}_{8,3}$ holds in general.

Lemma 5 *For every choice of representatives for the classes for t-subsets of $[n+t]$ there exists choices of representatives for the classes for t-multisets of $[n]$ so that the corresponding transition graphs $\mathcal{G}_{n+t,t}$ and $\mathcal{T}_{n,t}$ are isomorphic.*

Of course, the theorem holds in reverse as well, with the roles of $\mathcal{G}$ and $\mathcal{T}$ swapped, but we do not need that fact to prove Theorem 2.

Proof. Simply shift every digit of every class and corresponding representation down by one. Clearly $[f_1, \ldots, f_t]$ is a class for t-subsets of $[n+t]$ if and only if $[f_1 - 1, \ldots, f_t - 1]$ is a class for t-multisets of $[n]$. The same can be said for forms and for subsets/multisets. $\square$

Now we borrow the final fact from [2].

Fact 6 *Let $n_0(3) = 8$, $n_0(4) = 9$, and $n_0(6) = 17$. Then the transition graph $\mathcal{G}_{n,t}$ is Eulerian for $t \in \{3, 4, 6\}$ and $n \geq n_0(t)$ with $\gcd(n, t) = 1$.*

In light of Lemma 5 and the knowledge that $\gcd(n, t) = \gcd(n, n + t)$ we arrive at the following result.

Lemma 7 *Let $n_0(3) = n_0(4) = 5$ and $n_0(6) = 11$. Then the transition graph $\mathcal{T}_{n,t}$ is Eulerian for $t \in \{3, 4, 6\}$ and $n \geq n_0(t)$ with $\gcd(n, t) = 1$.*

The combination of Lemmas 7 and 4 yields Theorem 2, with the exception that $n_0(3) = 5$ instead of 4. However, a specific example for the case $t = 3$, $n = 4$ is given by the sequence S in (1) below. This concludes the proof of Theorem 2.

Note that this result is weaker than Conjecture 1 because the relative primality condition replaces the divisibility condition. The underlying mathematics for Lemma 7 comes from the result of [2] that no bad patterns exist if and only if $t \in \{3, 4, 6\}$ and $\gcd(n, t) = 1$.

3 Inductive Construction for $t = 3$

As noted in the introduction, one of the main results of this work is an inductive proof of theorem 2 in the case that $t = 3$. The proof is as follows.

For $t = 3$, the condition $n \,\middle|\, \binom{n+t-1}{t}$ implies that $n \equiv 1$ or $2 \pmod 3$. We will show that for $n \geq 4$, universal cycles on multisets exist whenever n satisfies $n \,\middle|\, \binom{n+t-1}{t}$, i.e. $n \,\middle|\, \binom{n+2}{3}$. We will prove this by induction on n as follows. We will start with a 3-Mcycle on $[n-3]$ of the form $stt'uv$, where $stt'uv$ is the concatenation of the substrings s, t, t', u, and v, where each of these strings is a substring over the alphabet $[n-3]$ with specific properties. From this string, we will construct a 3-Mcycle on $[n]$ of the form $STT'UV$, where $S = st$, $T = t'uv$, T' is a cyclic permutation of T, and U and V are to be described later.

Before describing the proof itself, we will define some terminology that will be useful for describing universal cycles. We say that a cyclic string $X = a_1a_2...a_k$ *contains* the multiset collection $\mathcal{I}$ if $\mathcal{I} = \big\{\{a_1, a_2, a_3\}, \{a_2, a_3, a_4\}, \ldots, \{a_{k-2}, a_{k-1}, a_k\}, \{a_{k-1}, a_k, a_1\}, \{a_k, a_1, a_2\}\big\}$, where each of these multisets must be distinct. Clearly $k = \binom{n+2}{3}$, since this is the number of 3-multisets on $[n]$.

For a string $X = a_1a_2...a_k$, we call the *lead-in* of X the substring a_1a_2 and the *lead-out* the substring $a_{k-1}a_k$.

Now, consider the collection of all 3-multisets over $[n]$. We shall partition this collection into four subcollections. Let $\mathcal{A}$ be the collection of all 3-multisets over $[n-3]$, and let $\mathcal{B}$ be the collection of all 3-multisets over $\{n-2, n-1, n\}$ and $[n-6]$ which contain at least one element from $\{n-2, n-1, n\}$. Let $\mathcal{C}$ be the collection of all 3-multisets with one or two elements from $\{n-5, n-4, n-3\}$ and one or two elements from $\{n-2, n-1, n\}$, and let $\mathcal{D}$ be the collection of all 3-multisets with one element from each of $[n-6]$, $\{n-5, n-4, n-3\}$, and $\{n-2, n-1, n\}$. We can see that $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ are disjoint, and that their union is the collection of all 3-multisets on $[n]$, as desired.

Now, let S be a 3-Mcycle on $[n-6]$, and since $1, 1, 1$ must occur somewhere in S and the beginning of S is arbitrary, we shall have S begin with $1, 1, 1$. We shall also select S so that its lead-out is $n-6, n-7$. Thus S, when considered as a cyclic string, contains all 3-multisets over $[n-6]$, and when considered as a non-cyclic string, contains all 3-multisets except $\{1, n-7, n-6\}$ and $\{1, 1, n-7\}$. Let T be a string over $[n-3]$ such that ST—the concatenation of S and T—is a 3–Mcycle over $[n-3]$. It is not clear that such a T must exist, but we shall find a specific example shortly. In the example we will find, T will begin with $1, 1$ and will end with $n-3, n-4$. Since T begins with $1, 1$, the string ST contains the multisets $\{1, n-7, n-6\}$, $\{1, 1, n-7\}$. We can see that the cyclic string ST contains all of the multisets in $\mathcal{A}$, and that when ST is considered as a non-cyclic string, it contains $\mathcal{A} \setminus \big\{\{1, n-4, n-3\},\ \{1, 1, n-4\}\big\}$. Now, consider the string T' obtained by taking T and replacing each instance of $n-5$ by $n-2$, $n-4$ by $n-1$, and $n-3$ by n. Since T contained all multisets over $[n-3]$ which contained at least one element from $\{n-5, n-4, n-3\}$, we have that T' contains all multisets over $\{n-2, n-1, n\}$ and $[n-6]$ which contain at least one element from $\{n-2, n-1, n\}$, i.e. T' contains all the multisets in $\mathcal{B}$. Since the lead-in of T is $1, 1$, the lead-in of T' is also $1, 1$, and since T ends with $n-3, n-4$, T' ends with $n, n-1$. If we consider the cyclic string STT', we can see that this string contains all the multisets in $\mathcal{A} \cup \mathcal{B}$, while the non-cyclic version of this string is missing the multisets $\{1, n-1, n\}$, $\{1, 1, n-1\}$.

For notational convenience, we will use the following assignments: $a := n-5$, $b := n-4$, $c := n-3$, $d := n-2$, $e := n-1$, and $f := n$. Now, we will construct the strings U and V. To do so, we shall consider the case where n is even and where n is odd. For n even, consider the following string:

$$\begin{aligned} V_e = \quad & \mathrm{be}(n-6)\mathrm{af}(n-7)\mathrm{be}(n-8)\mathrm{af}(n-9)...\mathrm{af1be} \\ & \quad \mathrm{ad}(n-6)\mathrm{ce}(n-7)\mathrm{ad}(n-8)\mathrm{ce}(n-9)...\mathrm{ce1ad} \\ & \qquad \mathrm{cf}(n-6)\mathrm{bd}(n-7)\mathrm{cf}(n-8)\mathrm{bd}(n-9)...\mathrm{bd1cfe}. \end{aligned}$$

We can see that this string contains every multiset in $\mathcal{D}$, as well as the multisets $\{a, b, e\}$, $\{a, d, e\}$, $\{a, c, d\}$, $\{c, d, f\}$, and $\{c, e, f\}$. Now, the following string (found with the aid of a computer) contains all of the multisets in $\mathcal{C} \setminus \big\{\{a, b, e\}, \{a, d, e\}, \{a, c, d\}, \{c, d, f\}, \{c, e, f\}\big\}$:

$$U_e = \mathrm{aaffc\ \ aeebb\ \ decec\ \ bddcc\ \ fbada\ \ dfbf}.$$

Note that while the multisets $\{b, b, f\}$ and $\{b, e, f\}$ are not present in the above string U_e, they are present in the concatenation of U_e with V_e. Similarly, while U_e does not contain $\{a, e, f\}$ and $\{a, a, e\}$, these multisets are present in the concatenation of T' with U_e.

Now, we can see that the string $STT'U_eV_e$ is a universal cycle over $[n]$ because the non-cyclic string STT' contained all the multisets in $\mathcal{A} \cup$

$\mathcal{B}\backslash\big\{\{1,n-1,n\},\ \{1,1,n-1\}\big\}$, and it is precisely the multisets $\{1,n-1,n\}$ and $\{1,1,n-1\}$ which are obtained by the wrap-around of the lead-out of V_e with the lead-in of S. The lead-in and lead-out of the other strings has been engineered so as to ensure that each multiset occurs precisely once.

Now, consider the case where n is odd. The corresponding strings V_o and U_o are

$$\begin{aligned} V_o = \quad & \text{be}(n-6)\text{af}(n-7)\text{be}(n-8)\text{af}(n-9)\text{...af2be} \\ & \quad \text{ad}(n-6)\text{ce}(n-7)\text{ad}(n-8)\text{ce}(n-9)\text{...ce2ad} \\ & \qquad \text{cf}(n-6)\text{bd}(n-7)\text{cf}(n-8)\text{bd}(n-9)\text{...bd2cfe.} \end{aligned}$$

and

$$U_o = \text{beb1f abd1c ffaae cbfbf dada1 eccfa eecdc dbd.}$$

The string V_o contains the same multisets as V_e, with the exception that V_o does not contain the nine multisets $\big\{\{1ad\},\{1ae\},...,\{1ce\},\{1cf\}\big\}$, and the string U_o contains the same multisets as U_e, with the exception that it contains the additional nine multisets listed above. The concatenation of V_o and U_o with the other strings works the same way as their even counterparts.

This completes the induction proof, since the string ST is a 3-Mcycle over $[n-3]$ (taking the place of S in the previous iteration of the induction), and the string $T'UV$ extends this 3-Mcycle to $[n]$ (taking the place of T in the previous iteration of the induction). Also note that $T'UV$ begins with $1,1$ and ends with $n,n-1$, as required for the induction hypothesis.

Thus, all that remains is the find a base case from which the induction can proceed. A possible base case (there are many) for $n=10$ is

$$\begin{aligned} S &= 11144\ 42223\ 33121\ 24343 && (1) \\ T &= 11522\ 63374\ 45166\ 27732\ 57366\ 77135\ 34641\ 71555\ 36127\ 42556 \\ &\quad\ 66477\ 75526\ 4576, \end{aligned}$$

which would lead to

$$\begin{aligned} T' &= 11822\ 93304\ 48199\ 20032\ 80399\ 00138\ 34941\ 01888\ 39120\ 42889 \\ &\quad\ 99400\ 08829\ 4809 \\ U &= 55007\ 59966\ 89797\ 68877\ 06585\ 8060 \\ V &= 69450\ 36925\ 01695\ 84793\ 58279\ 15870\ 46837\ 02681\ 709, \end{aligned}$$

Where "0" denotes 10 and the spacings have been added to increase readability.

A possible base case for $n = 11$ is

$$\begin{aligned} S &= 11122\ 23114\ 22513\ 32444\ 33352\ 54541\ 43555 \\ T &= 11657\ 43822\ 74468\ 54661\ 72736\ 18157\ 31888\ 77556\ 6688\ 57262 \\ &\quad 58536\ 21848\ 47776\ 41773\ 38826\ 67836\ 36428\ 7, \end{aligned}$$

The corresponding strings T', U, and V can be found using the method outlined above.

Thus we have established a method to generate universal cycles on multisets for $t = 3$, $n \equiv 1, 2 \pmod 3$.

4 Conversion Construction for $t = 3$

In this proof, we construct a 3-Mcycle by modifying a 3-Ucycle. (We know from [2] that 3-Ucycles on $[n]$ exist for all $n \geq 8$ not divisible by 3.) Before giving the proof, we introduce two terms. We call each element of $[n]$ a *letter*, and each a_i in the Ucycle $X = a_1 \dots a_k$ a *character*. To summarize, a 3-Mcycle on $[n]$ is made up of $\binom{n+t-1}{t}$ characters, each of which equals one of n letters.

To demonstrate the proof's technique, we will first use an argument similar to it to create 2-Mcycles from 2-Ucycles. We start with this 2-Ucycle on [5]:

$$1234513524.$$

Then, we repeat the first instance of every letter to create the following 2-Mcycle:

$$112233445513524.$$

The technique works because repeating a character a_i as above adds the multiset $\{a_i, a_i\}$ to the Ucycle and has no other effect.

To use this technique on 3-Ucycles, we repeat not single characters, but pairs of characters. For example, changing

$$\dots a_{i-1}a_ia_{i+1}a_{i+2}\dots$$

to

$$\dots a_{i-1}a_ia_{i+1}a_ia_{i+1}a_{i+2}\dots$$

has only the effect of adding the 3-multisets $\{a_i, a_i, a_{i+1}\}$ and $\{a_i, a_{i+1}, a_{i+1}\}$ to the cycle. In order to use this technique, we will need to know which consecutive pairs of letters appear in a 3-Ucycle. For instance, the following 3-Ucycle (generated using methods from [2]) on [8] contains every unordered pair of letters as consecutive characters but $\{1, 5\}$, $\{2, 6\}$, $\{3, 7\}$, and $\{4, 8\}$:

1235783 6782458 3457125 8124672 5671347 2346814 7813561 4568236,

where spaces in the cycle are added to increase readability. This Ucycle is missing 4 pairs, which happens to be $n/2$. This is no coincidence: in fact, this is the most number of pairs that a 3-Ucycle can fail to contain.

Lemma 8 *No two unordered pairs not appearing as consecutive characters in a 3-Ucycle have a letter in common. A 3-Ucycle can hence be missing at most $n/2$ pairs of letters.*

Proof. Suppose that we have a 3-Ucycle that contains neither a and b as consecutive characters, nor a and c as consecutive characters, where $a, b, c \in [n]$. Then the 3-Ucycle does not contain the 3-subset abc, for all permutations of abc contain either a and b consecutively, or a and c consecutively. But this is a contradiction, as a 3-Ucycle by definition contains all 3-subsets.

Hence, no two pairs of characters missing in the 3-Ucycle can have a letter in common. By the pigeonhole principle, the 3-Ucycle can be missing at most $n/2$ pairs of letters.

With this lemma, we can finish our proof, creating a 3-Mcycle on $[n]$ whenever n is not divisible by 3. Let X be a 3-Ucycle on $[n]$. Let $x_1, \ldots, x_n$ be a permutation of $[n]$ such that

- x_1 equals the first character in X.
- x_n equals the last character in X.
- If x is even, the list $\{x_1, x_2\}, \{x_3, x_4\}, \ldots, \{x_{n-1}x_n\}$ contains all unordered pairs of letters not contained as consecutive characters in X, which is possible by our lemma. (If X is missing exactly $n/2$ pairs of letters, these pairs will be exactly the pairs missing from X. If X is missing fewer than $n/2$ pairs of letters, then the pairs consist of all missing pairs of letters, plus the remaining letters paired arbitrarily.)

  If x is odd, one of the following lists contains all unordered pairs of letters not contained as consecutive characters in X:

1. $\{x_1, x_2\}, \ldots, \{x_{n-2}, x_{n-1}\}$
2. $\{x_1, x_2\}, \{x_4, x_5\}, \{x_6, x_7\} \ldots, \{x_{n-1}, x_n\}$
3. $\{x_2, x_3\}, \ldots, \{x_{n-1}, x_n\}$

This is possible by our lemma. There can be at most $(n-1)/2$ missing pairs of letters in X, and depending on whether x_1, x_n, or both is a member of a missing pair, one of the above lists can contain all the missing pairs. (As in the even case, it does not present any problems if X is missing fewer than $(n-1)/2$ pairs.)

Make X' by repeating the first instance of every unordered pair of letters in X except for $\{x_1, x_2\}, \{x_2, x_3\}, \ldots, \{x_{n-1}, x_n\}, \{x_n, x_1\}$. The cycle X' now contains all multisets except

$$\{x_1, x_1, x_1\}, \ldots, \{x_n, x_n, x_n\}$$

$\{x_1, x_1, x_2\}, \{x_1, x_2, x_2\}, \{x_2, x_2, x_3\}, \{x_2, x_3, x_3\}, \ldots, \{x_n, x_n, x_1\}, \{x_n, x_1, x_1\}$.

Now, add the string $x_1x_1x_1x_2x_2x_2 \ldots x_nx_nx_n$ to the end of X' to create X''. This provides exactly the missing multisets, creating a 3-Mcycle.

For example, when $n = 8$, we start with the following 3-Ucycle:

$$\begin{aligned} X = &1235783\ 6782458\ 3457125\ 8124672 \\ &5671347\ 2346814\ 7813561\ 4568236. \end{aligned}$$

The 3-Ucycle X does not contain the pairs $\{1, 5\}$, $\{2, 6\}$, $\{3, 7\}$, and $\{4, 8\}$. Hence, we set

$$\begin{aligned} x_1 = 1,\ x_2 = 5,\ x_3 = 3,\ x_4 = 7 \\ x_5 = 4,\ x_6 = 8,\ x_7 = 2,\ x_8 = 6. \end{aligned}$$

Note that x_1 equals the first character of X, and x_8 equals the last.

Now, we repeat the first instance of every unordered pair except for $\{1, 5\}$, $\{5, 3\}$, $\{3, 7\}$, $\{7, 4\}$, $\{4, 8\}$, $\{8, 2\}$, $\{2, 6\}$, and $\{6, 1\}$. (Note that four of these pairs do not appear in X. If some of these pairs actually did appear in X, because X was missing fewer than $n/2$ pairs of letters, it would not affect the proof.)

$$\begin{aligned} X' = &12123235757878383\ 63676782424545858\ 3434571712525\ 81812464672 \\ &56567131347\ 2723468681414\ 7813561\ 4568236. \end{aligned}$$

Finally, we add the string $x_1x_1x_1 \dots x_nx_nx_n$ to complete the Mcycle.

$$X'' = 12123235757878383\ 636767824245458 58\ 3434571712525\ 81812464672$$
$$56567131347\ 2723468681414\ 7813561\ 4568236$$
$$111555333777444888222666.$$

5 Remarks

The proofs in Sections 3 and 4 suggest natural extensions to the $t = 4$ and larger cases. Moreover, they may prove useful by their introduction of new techniques for approaching Ucycles. The section 3 proof is notable for its use of induction, a technique which has not been used to create Ucycles. This is especially promising in light of the many potential base cases provided by Jackson [3] for $t \leq 11$. The section 4 proof, while it is tied to Ucycles, is not tied to any particular approach for creating Ucycles, which is not true of the technique in Section 2. Since the necessary condition for the existence of Ucycles for t-subsets of $[n]$ is that n divides $\binom{n}{t}$, and since $\frac{1}{n}\binom{n+t-1}{t} = \frac{1}{n+t}\binom{n+t}{t}$, we see that the condition for t-Mcycles of $[n]$ is the same as the condition for t-Ucycles of $[n + t]$. Thus it is reasonable to assume that some sort of transformation between the two exists.

For values of n and t for which Mcycles do exist, one interesting question is how many Mcycles exist. Clearly each Mcycles has $n!$ representations, since there are $n!$ permutations of $1, \dots, n$. However, when searching for Mcycles using a computer, vast numbers of *distinct* (i.e. not differing merely by a permutation of $1, \dots, n$) Mcycles were found. Currently, it is not clear whether $N(n, t)$, the number of distinct Mcycles for a given value of n and t, is a function that can be approximated well.

References

[1] F.R.K. Chung, P. Diaconis, R.I. Graham, *Universal Cycles for Combinatorial Structures*, Discrete Math. **110** (1992), 43–59.

[2] G. Hurlbert, *On Universal Cycles for k-subsets of an n-set*, SIAM J. Discrete Math. **7** (1994), 598–605.

[3] B. Jackson, personal communication (2004).

[4] D.E. Knuth, The Art of Computer Programming, Vol. 4, Fasc. 3: Generating All Combinations and Permutations. Addison-Wesley (2005).