Tables of Assembly Graphs

November 10, 2009

1 Introduction

This file is located at http://www.math.usf.edu/∼saito/DNAweb/table.pdf, and is one of the files posted at the web site http://www.math.usf.edu/∼saito/DNAweb/. The web site contains information and results on *assembly graphs*, that are mathematical models for DNA recombination processes. See http://www.math.usf.edu/∼saito/DNAweb/bachground.pdf for background materials, which will be referred to as [BG].

This file will be updated often during 2009–2013 to post our current results on making tables of assembly graphs.

2 Assembly graphs with a single transverse component and two end points

Recall from [BG] that an assembly graph being *simple* means having a single transverse component, and *irreducible* means that it is not decomposed into two non-trivial assembly graphs attached at end points. In this case, assembly graphs are represented by double occurrence words, where we choose positive integers as an alphabet. There are no a priori orientations so that the word corresponding to the opposite orientation is considered equivalent to the word corresponding to the original orientation. Choose an arbitrary orientation, and trace the graph from the initial point to the terminal point. For convenience we choose 1 as the first letter assigned to the first rigid 4-vertex that is encountered, 2 as the second letter, unless the second letter is also 1 (in this case a small loop brings us back to the first vertex), and so on.

2.1 Length 1 and 2

It is easy to see that simple irreducible assembly graphs of length 1 is represented by 11, length 2 are by 1212 and 1221. They are depicted in Fig. 1 from left to right. All these have the assembly number 1.

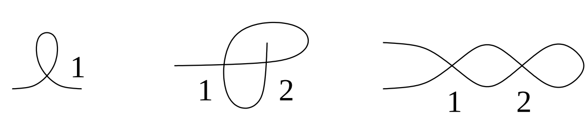

Figure 1: List for lengths 1 and 2

2.2 Length 3

For simple, irreducible assembly graphs with two end points, we have the following.

Lemma 2.1 *Simple irreducible assembly graphs of length* 3 *consist of those represented by* 121323*,* 121332*,* 122331*,* 123123*,* 123132*,* 123231*,* 123312*,* 123321*.*

Proof. We may assume that a word starts with 12, since 11 would give a reducible word.

If it starts with 121, then from irreducibility, it must continue with 3, so that it starts with 1213. There are two possibilities in this case: 121323 and 121332.

If it starts with 122, again it must continue with 3, such as 1223. The word 122313 is equivalent to 121332, and we have only 122331, in this case. If it starts with 123, among 6 possibilities for the remaining three letters, 123213 is equivalent to 123132, and others are distinct. □

These are depicted in Fig. 2. All these have the assembly number 1.

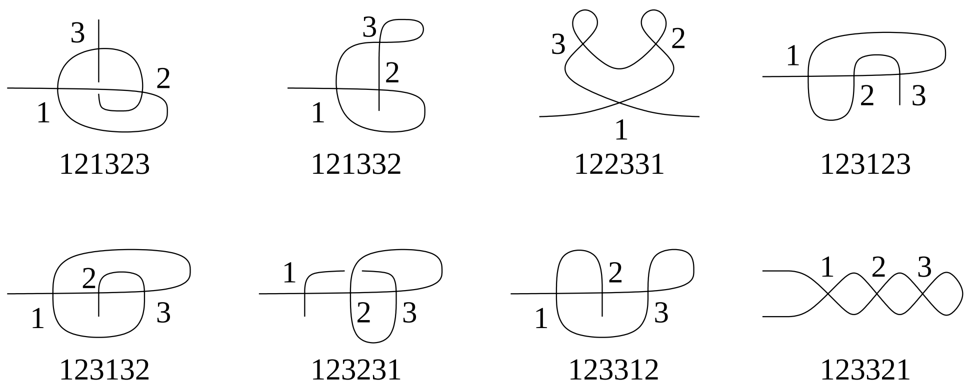

Figure 2: List for length 3

3 Connected assembly graphs with two transverse components and two end points

In this section we consider assembly graphs of two transverse components and with two end points. In this case, one of the two transverse component has two end points (an arc), and the other component is circular. We use the notation $[a_1 \cdots a_m; b_1 \cdots b_n]$ for assembly words, where $a_1 \cdots a_m$ represents the arc component, and $b_1 \cdots b_n$ represents the circular component. More generally, words for an assembly graph of m arc components and n circular components are represented by $[u_1, \dots, u_n; v_1, \dots, v_n]$, where $\{u_i \mid i = 1, \dots, m\}$ (resp. $\{v_j \mid j = 1, \dots, n\}$ are unordered words, and each word u_i (resp. v_j) is defined up to reverse (resp. circlic permutations and reverse) and renaming of letters.

Definition 3.1 An assembly word $[u; v_1, \dots, v_k]$ of $k+1$ transverse components and two end points is *reducible* if there are assembly words $[u_1; x_1, \dots, x_m]$ and $[u_2; y_1, \dots, y_n]$ of $m+1$ and $n+1$ transverse components, respectively, and two end points, such that $[u; v_1, \dots, v_k] = [u_1 \circ u_2; x_1, \dots, x_m, y_1, \dots, y_n]$, where we assume that letters in $\{x_i\}_{i=1}^m$ and $\{y_j\}_{j=1}^n$ are disjoint, and $\{v_1, \dots, v_k\} = \{x_1, \dots, x_m, y_1, \dots, y_n$ (in particular, $k = m+n$).

3.1 Length 1 and 2

It is easy to see that irreducible assembly graphs of length 1 is represented by $[1;1]$, length 2 are by $[1;122]$, $[12;12]$ and $[121;2]$. They are depicted in Fig. 3 from left to right. All these have the assembly number 1.

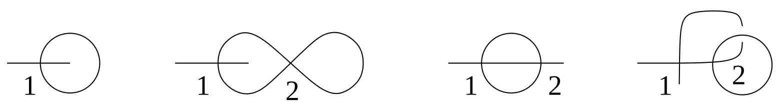

Figure 3: List for lengths 1 and 2

3.2 Length 3

For length 3 we obtain the following list.

Lemma 3.2 (M. A. Nikiforou) *Irreducible connected assembly graphs of length* 3 *with two transverse components and two end points consist of those represented by* $[1;12233]$, $[1;12332]$, $[1;12323]$, $[12;1233]$, $[121;233]$, $[123;123]$, $[1213;23]$, $[1223;13]$, $[1231;23]$, $[1232;13]$, $[12132;3]$, $[12231;3]$, $[12313;2]$, $[12321;3]$.

Proof. If the arc component is the word 1, then the circular component has one copy of the letter 1, and by cyclic permutations we may assume that the word for the circular component starts with 1. Furthermore, we may assume as before that it starts with 12, from the convention that we follow the numerical order to name the vertices. Then the choices for the circular component are 12233, 12323 and 12332.

If the word for the arc component is 12, and the word for the circular component starts with 12, then the circular component has to be 1233.

We examine the cases where the arc components are 3-letter words. If the word is 121, then the circular component is forced to be 233. The word 122 for the arc component gives a reducible graph, so it is not listed. The word 123 for the arc component forces the circular component to be 123 as well, up to equivalence.

Next we examine the cases when the arc component has 4 letters. If it starts with 121, then from connectivity it must continue with 3, 1213, and we obtain $[1213, 23]$. If it starts with 122, again it must continue with 3, 1223, and we obtain $[1223, 13]$. If it starts with 123, then there are three possibilities: 1231, 1232, and 1233, the last of which is reducible. Hence we obtain $[1231, 23]$ and $[1232, 13]$, in this case.

Finally, we look at the cases where the arc components are 5-letter words: Again we may assume that a word starts with 12, since 11 will give a reducible word. If it starts with 121, then it must be 12123 or 12132, since 12133 is reducible. We obtain $[12123, 3]$ (reducible), and $[12132, 3]$. If it starts with 122, then it must be 12213 (which is reducible) or 12231, since 12233 is again reducible. If it starts with 123 then there are six possibilities. We have $[12312, 3]$, $[12313, 2]$ ($\cong [12132, 3]$), $[12321, 3]$, $[12323, 1]$ (reducible), $[12331, 2]$ ($\cong [12231, 3]$), $[12332, 1]$ ($\cong [12213, 3]$, which is reducible). □

These are depicted in Fig. 4.

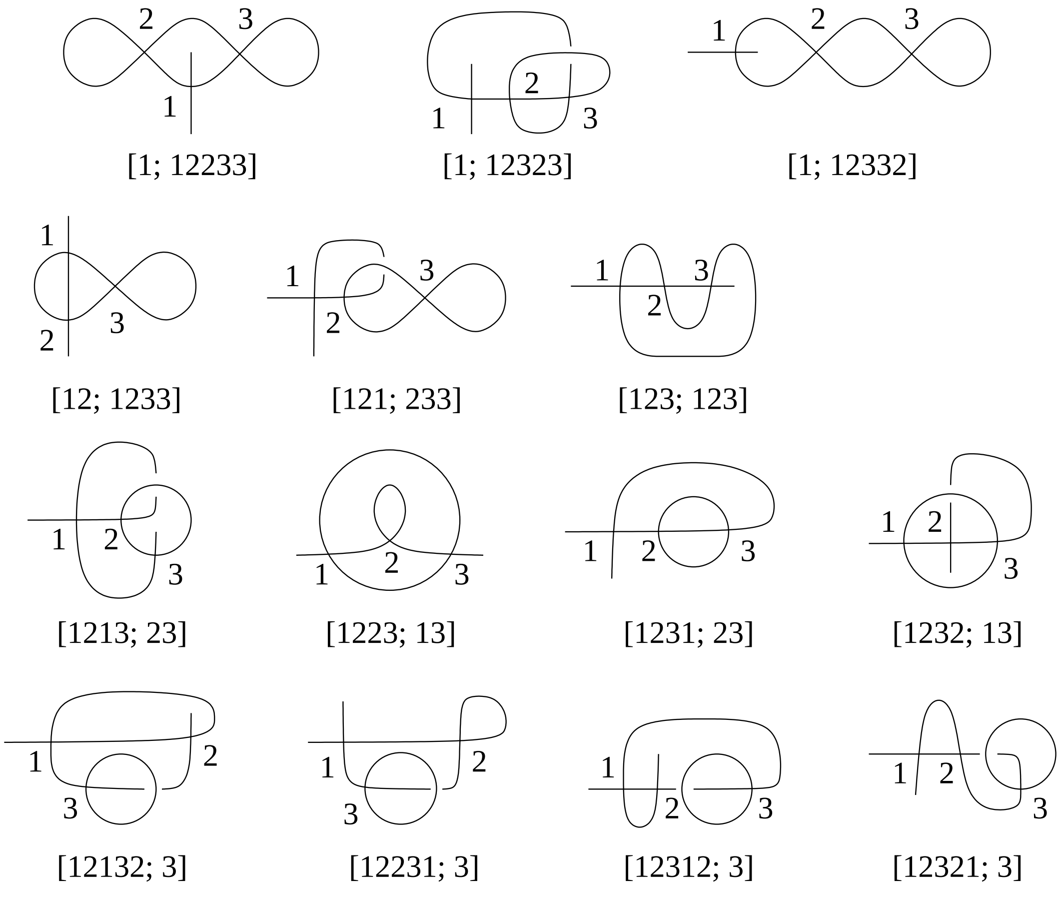

Figure 4: List for lengths 3