

CODES OF ALL SATURATED ACYCLIC ALKANE ISOMERS CONTAINING FOURTEEN CARBON ATOMS

István Lukovits

Chemical Research Center, H-1525 Budapest, P. O. B. 17, Hungary

(Received August 13, 2002)

ABSTRACT. Codes representing the structural formulas of all 1858 acyclic saturated alkane isomers containing fourteen carbon atoms have been generated and listed. From each code the underlying adjacency matrix can be restored by using a simple procedure. The procedures used to select the structures from a greater set of labelled structures were briefly discussed.

INTRODUCTION

The enumeration of alkane isomers (including optical isomers) was accomplished already in the thirties of the last century [1]. Despite of the fact that this (mathematical) problem was solved, chemists were interested in it, since they wanted to possess an exhaustive (i.e. complete) list of non-redundant structures. This problem was also solved. The approach advocated by Knop et al [2] was named the N-tuple code method. N-tuple codes are based on physical trees (see next section), which is a graph theoretical concept. (Knop et al. [2] published the formulas of all isomers of alkanes containing 12 carbon atoms.) It seems that the N-tuple approach is one of the most successful methods, which is still being extended [3]. Another approach based on the “lowest-degrees-first” (LDF) method [4,5], was proposed twenty years after the paper by Knop et al. appeared. By using the LDF method all codes, representing structural formulas of acyclic alkane isomers containing 13 carbon atoms were generated [6]. The aim of this paper was to generate and list all 1858 codes of (acyclic) alkane isomers, containing 14 carbons. From these codes the underlying adjacency matrices and (ultimately) the respective structural formulas can be restored.

THE OPTIMAL NUMBERING OF VERTICES

Chemical (structural) formulas are in fact colored graphs. A graph G is composed of a set of N vertices (these correspond to the atoms) and a set of edges connecting vertices. Edges represent chemical bonds. In acyclic graphs the number of edges is equal to $N - 1$. In what follows we shall consider non-colored graphs, only, as hydrogen depleted structural formulas of alkanes can be colored by using a single color. Computer manipulation of structural formulas is most easily accomplished by using the adjacency matrix \mathbf{A} . \mathbf{A} is an $N \times N$ symmetric matrix, and $A_{i,j} = 1$ if vertices i and j ($1 \leq i, j \leq N$) are connected by a (single) bond,

and $A_{i,j} = 0$, otherwise. Note that $A_{i,i}$ (the diagonal element) is also zero for any vertex i . Any formula can be represented by its corresponding adjacency matrix, provided the atoms (vertices) have already been numbered and all formulas can be restored from its underlying adjacency matrix. Numbering of vertices is a crucial step, and (as we shall see) if the numbers are allocated to vertices properly, then the time needed for the generation procedure can be reduced substantially. Expression 'structure' denotes hereafter a hydrogen suppressed chemical formula with numbered (from 1 through N) vertices.

Matrix **A** is symmetric, therefore it is enough to consider the upper right triangular part of it. Usually, if the numbering is done in a unsystematic way, this upper triangular part will contain several non-zero entries in single column. However, if the numbering is done in a systematic way, we may obtain adjacency matrices in which the upper triangular part contains one, and only one non-zero entry in each column. The simplest method is through 'physical trees' [2]. A structure is a physical tree if after assigning number 1 to any vertex, the following ordinals are only attached to vertices being adjacent to an already numbered vertex. Figure 1 shows a physical tree containing fourteen carbon atoms. Each vertex, except vertex 1, has only a single neighbour with a lower ordinal. The reader can verify by working out the adjacency matrix, that in each column of the upper triangle there is only a single entry equal to one.

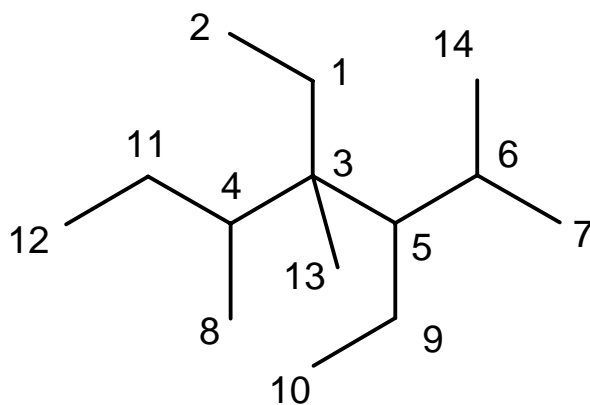


Figure 1. A physical tree containing fourteen carbon atoms.

A Morgan tree [6] is a physical tree in which numbering (except vertex 1) starts with a vertex adjacent the vertex with the lowest ordinal and still possessing not-numbered neighbours. Figure 2 illustrates the numbering of vertices in a Morgan tree.

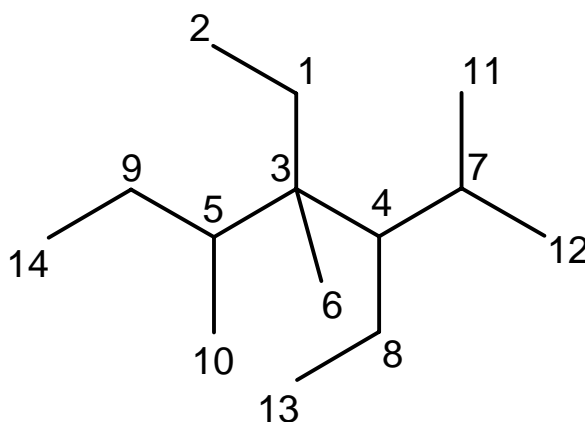


Figure 2. A Morgan tree containing fourteen carbon atoms.

Again by working out the adjacency matrix of the structure depicted in Figure 2 we obtain Figure 3. Observe the ‘stairs-like’ pattern of the non-zero entries, which accounts for the fact, why the number of Morgan trees is much less than the number of physical trees (Figure 3).

Since the upper right triangle of physical (and Morgan) trees contains a single non-zero entry, the adjacency matrix can be compressed, giving rise to the compressed adjacency matrix (CAM). The CAM is a vector \mathbf{C} containing $N - 1$ entries and

$$\text{If } A_{i,k+1} = 1 \text{ then } C_k = i \quad (1)$$

The CAM corresponding to the adjacency matrix depicted in Figure 2 is:

$$\mathbf{C}_{\text{Morgan}} = (1,1,3,3,3,4,4,5,5,7,7,8,9) \quad (2)$$

By using eq. 1 the underlying adjacency matrix (and therefore the underlying structure, too) can be restored from \mathbf{C} . Examples how to restore the underlying structure from a Morgan tree without using the adjacency matrix representation, have been given earlier [6,7].

A ‘lowest-degree-first’ (LDF) tree is a Morgan tree (and therefore it is also a physical tree) in which the numbering starts with the vertex of the lowest degree. (The degree d_i of vertex i denotes the number of edges incident at i .) As a consequence of this regulation, vertex 1 is always an endpoint. If the order can not be determined by inspecting the first neighbours, then the second order neighbours should be inspected. If these are also not decisive, then the third order neighbours should be examined, etc.

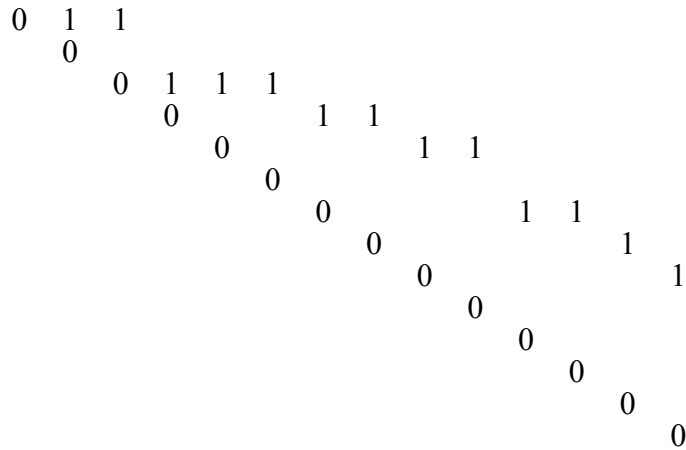


Figure 3. Upper right triangular of the adjacency matrix corresponding to the structure depicted in Figure 2. In order to expose the ‘stairs-like’ pattern, all zeroes, except those in the diagonal, have been omitted.

The CAM corresponding to the LDF structure (Figure 4) is:

$$\mathbf{C}_{\text{LDF}} = (1,2,3,3,5,5,5,7,8,8,10,11,11) \quad (3)$$

It was proved [4] that for any structure $\mathbf{C}_{\text{LDF}} \geq \mathbf{C}_{\text{Morgan}}$, i.e. codes of LDF structures are maximal and therefore LDF structures are unique. In our example $\mathbf{C}_{\text{LDF}} = (1,2,3,3,5,5,5,7,8,8,10,11,11) > (1,1,3,3,3,4,4,5,5,7,7,8,9) = \mathbf{C}_{\text{Morgan}}$, because the second entry in \mathbf{C}_{LDF} is greater than the same entry in $\mathbf{C}_{\text{Morgan}}$. (The first elements are identical.) Note that numbers related to endpoints - except vertex 1 - do not appear in CAMs.

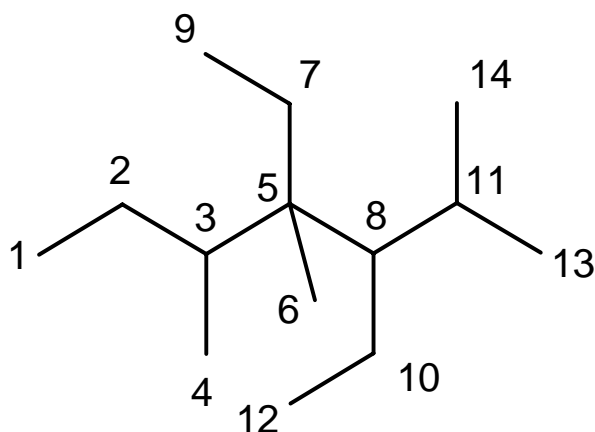


Figure 4. An LDF tree.

Since the number of Morgan trees comprise only a small fraction of the number of physical trees [5], and each tree can be numbered such that it is a Morgan tree, the exhaustive and non-redundant generation of all structural formulas corresponding to isomers of acyclic alkanes is accomplished via the generation of codes of Morgan trees [4,5]. It has to be noted that significant portions of Morgan codes do not contain maximal codes and therefore need not be inspected and generated [4]. Once the codes (CAMs) of Morgan trees have been obtained, various rules (see next section) can be used to eliminate non-LDF (and therefore non-maximal) codes.

ELIMINATION OF REDUNDANT STRUCTURES

Generation of codes of Morgan trees starts with the CAM of the lowest lexicographic value (1,2,2,2,...,2), and then CAMs of increasing lexicographic values are generated in a systematic way. Note that each vertex i (except vertex 1) appears $d_i - 1$ times in the CAM. Several portions of CAMs may not be generated because they do not contain maximal (LDF) codes [4]. Once the codes have been generated, they are screened by using the rules listed below. The rules are ordered in terms of increasing complexity and all together are sufficient to select LDF codes from the greater set of Morgan codes. The first two rules (both are done in linear time) are used for screening of the codes, since they limit the maximal degree and also make sure that codes violating the LDF principle in trivial manner, are not processed further. In order to explain the rules let us introduce the concept of the valence code and the path code of a tree. The proofs of these rules were presented in ref [7].

The valence code \mathbf{v} is an N - vector containing the degrees of the vertices. As an example consider the structure depicted in Figure 5: $\mathbf{v} = (1-2-4-1-3-3-2-3-1-2-1-1-1-1)$. Valence codes related to different branches are also possible, as an example consider two paths starting at vertex 1 in Figure 5. For example the valence codes of paths 1-9 and 1-11 are $\mathbf{v}_{1-9} = (1-2-4-3-1-0)$, and $\mathbf{v}_{1-11} = (1-2-4-3-2-1)$, respectively. The entry equal to zero in \mathbf{v}_{1-9} has been inserted in order to set the two vectors on an equal footing.

The path code \mathbf{P}_1 is the set of all possible valence codes of paths starting at vertex 1. Path code \mathbf{P}_2 is the set of all possible valence codes starting at another specified endpoint. Both \mathbf{P}_1 and \mathbf{P}_2 contain the valence codes in lexicographic order.

Rule 1. Any structure containing at least one vertex of degree greater than four must be deleted. Therefore any CAM, in which the same entry appears four or more times, must be deleted.

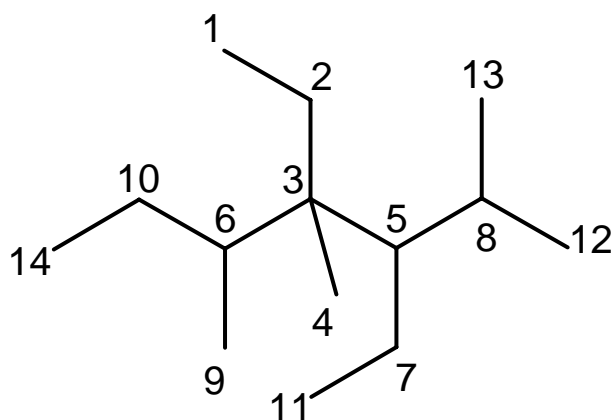


Figure 5. Another Morgan tree containing fourteen carbon atoms.

Rule 2. Any CAM with $C_i = C_{i+1}$ and $d_{i+1} = d_{i+2}$ must be deleted [5,6], since in this case the ‘lowest degrees first’ rule has been violated. As an example consider the structure depicted in Figure 2, the CAM is: $\mathbf{C} = (1,1,3,3,3,4,4,5,5,7,7,8,9)$. Entries 6 and 7 are identical (4), but vertex 7 has a greater degree than vertex 8 violating the LDF principle.

Rule 3. Compare two paths connecting vertex 1 with vertex m and n , where m and n are endpoints. Let’s assume, that the first non-common vertex in path 1- m is k , and is k' in path 1- n . Path 1- m precedes path 1- n if $k < k'$. Delete structure if $\mathbf{v}_{1-m} > \mathbf{v}_{1-n}$ Example (Figure 5): branch 1-11 precedes branch 1-9 because the first non-common vertex is 5 in the former and 6 in the latter. On the other hand $\mathbf{v}_{1-9} > \mathbf{v}_{1-11}$ (see preceding paragraph), indicating that the actual numbering violates the LDF principle. Rule 3 is used to detect all incorrectly labelled branches, assuming that vertex 1 has been selected correctly.

Rule 4. Let \mathbf{P}_1 and \mathbf{P}_2 , be two path codes. Delete structure if $\mathbf{P}_1 > \mathbf{P}_2$. Example: consider Figure 5 and find the path codes (i.e. six valence codes) with respect to vertex 1 and 14:

\mathbf{P}_1	\mathbf{P}_2 ,
1-2-4-1-0-0-0	1-2-3-1-0-0-0
1-2-4-3-1-0-0	1-2-3-4-1-0-0
1-2-4-3-2-1-0	1-2-3-4-2-1-0
1-2-4-3-2-1-0	1-2-3-4-3-2-1
1-2-4-3-3-1-0	1-2-3-4-3-3-1
1-2-4-3-3-1-0	1-2-3-4-3-3-1

where the zeroes have been added to set the path codes on an equal footing. Already the first lines in \mathbf{P}_1 and \mathbf{P}_2 are different and therefore $\mathbf{P}_1 > \mathbf{P}_2$, and the respective numbering (Figure 5) is incorrect. Rule 4 ensures that vertex 1 is placed optimally (i.e. to yield the maximal CAM), and avoids the time-consuming renumbering of a structure.

The actual calculations were done by using a computer program performing the generation and selection procedures [5,6].

3 3 5 6 6 7 8 8 8 12 12
3 3 5 6 6 7 8 8 10 11 11
3 3 5 6 6 7 8 10 10 10 13
3 3 5 6 6 8 8 8 9 10 11
3 3 5 6 6 8 8 9 9 10 10
3 3 5 6 6 8 8 10 10 11 12
3 3 5 6 6 8 9 9 9 11 12
3 3 5 6 6 8 9 9 11 12 12
3 3 5 6 6 8 9 10 11 12 12
3 3 5 6 6 8 9 10 11 12 12
3 3 5 6 7 7 7 9 10 12 12
3 3 5 6 7 7 7 10 11 11 13
3 3 5 6 7 7 8 9 9 10 10
3 3 5 6 7 7 8 9 11 12 12
3 3 5 6 7 7 8 9 11 12 12
3 3 5 6 7 7 9 9 10 10 13
3 3 5 6 7 7 9 10 11 11 11
3 3 5 6 7 7 9 10 11 11 11
3 3 5 6 7 8 8 8 11 11 11
3 3 5 6 7 8 8 9 10 10 13
3 3 5 6 7 8 8 10 11 11 11
3 3 5 6 7 8 9 9 10 11 11
3 3 5 6 7 8 9 10 10 11 12
3 4 4 4 5 5 6 6 7 7 7
3 4 4 4 5 6 6 7 7 7 13
3 4 4 4 5 6 6 7 7 12 12
3 4 4 4 5 6 7 7 7 12 12
3 4 4 4 5 6 7 7 9 9 13
3 4 4 4 5 6 7 7 11 11 11
3 4 4 4 5 6 7 9 10 10 10
3 4 4 4 5 6 7 10 10 12 12
3 4 4 4 6 6 7 7 7 8 9
3 4 4 4 6 6 7 7 9 10 11
3 4 4 4 6 6 7 7 11 12 12
3 4 4 4 6 6 7 7 8 8 8
3 4 4 4 6 7 7 7 8 11 11
3 4 4 4 6 7 7 7 8 12 12
3 4 4 4 6 7 7 7 11 11 13
3 4 4 4 6 7 7 8 8 9 10
3 4 4 4 6 7 7 8 9 10 10
3 4 4 4 6 7 7 8 11 11 13
3 4 4 4 6 7 7 9 10 10 13
3 4 4 4 6 7 7 9 10 10 13
3 4 4 4 6 7 7 10 11 11 11
3 4 4 4 6 7 8 9 9 9 13
3 4 4 4 6 7 8 9 11 11 13
3 4 4 4 6 7 9 9 11 11 11
3 4 4 4 6 7 9 10 10 12 12
3 4 4 4 6 7 9 10 10 13
3 4 4 4 7 7 7 8 9 9 13
3 4 4 4 7 7 8 9 11 11 13
3 4 4 4 7 7 9 9 10 11 11
3 4 4 4 7 7 9 10 10 11 12

3 3 5 6 6 7 8 8 9 9 9
3 3 5 6 6 7 8 8 11 11 11
3 3 5 6 6 7 8 10 10 11 12
3 3 5 6 6 8 8 8 10 11 11
3 3 5 6 6 8 8 9 10 10 10
3 3 5 6 6 8 8 10 10 12 12
3 3 5 6 6 8 9 9 9 12 12
3 3 5 6 6 8 9 10 10 10 13
3 3 5 6 7 7 7 8 9 10 10
3 3 5 6 7 7 7 10 10 10 13
3 3 5 6 7 7 7 10 11 12 12
3 3 5 6 7 7 8 9 9 11 12
3 3 5 6 7 7 8 9 9 11 12
3 3 5 6 7 7 9 9 11 12 12
3 3 5 6 7 7 9 10 11 11 13
3 3 5 6 7 8 8 8 11 11 13
3 3 5 6 7 8 8 9 10 12 12
3 3 5 6 7 8 8 10 11 11 13
3 3 5 6 7 8 8 9 10 11 11
3 3 5 6 7 8 9 10 10 12 12
3 4 4 4 5 5 6 6 7 7 13
3 4 4 4 5 6 6 7 7 8 8
3 4 4 4 5 6 7 7 7 8 9
3 4 4 4 5 6 7 7 8 9 9
3 4 4 4 5 6 7 7 9 10 11
3 4 4 4 5 6 7 7 11 11 13
3 4 4 4 5 6 7 9 10 10 13
3 4 4 4 5 6 7 10 11 11 11
3 4 4 4 6 6 7 7 7 9 9
3 4 4 4 6 6 7 7 9 11 11
3 4 4 4 6 7 7 7 8 8 8
3 4 4 4 6 7 7 7 8 12 12
3 4 4 4 6 7 7 7 11 11 13
3 4 4 4 6 7 7 8 8 10 10
3 4 4 4 6 7 7 8 10 10 10
3 4 4 4 6 7 7 8 11 11 13
3 4 4 4 6 7 7 9 10 12 12
3 4 4 4 6 7 7 10 11 11 13
3 4 4 4 6 7 8 9 9 10 10
3 4 4 4 6 7 8 9 11 12 12
3 4 4 4 6 7 9 9 11 11 13
3 4 4 4 6 7 9 10 11 11 11
3 4 4 4 7 7 7 9 9 10 10
3 4 4 4 7 7 7 9 10 12 12
3 4 4 4 7 7 7 10 11 11 13
3 4 4 4 7 7 8 9 9 10 10
3 4 4 4 7 7 8 9 11 11 13
3 4 4 4 7 7 9 9 11 11 11
3 4 4 4 7 7 9 10 10 12 12

3 3 5 6 6 7 8 8 9 9 11
3 3 5 6 6 7 8 8 11 11 13
3 3 5 6 6 7 8 10 10 12 12
3 3 5 6 6 8 8 11 11 11
3 3 5 6 6 8 8 9 10 10 13
3 3 5 6 6 8 8 10 11 11 11
3 3 5 6 6 8 9 9 10 11 11
3 3 5 6 6 8 9 10 10 11 12
3 3 5 6 7 7 7 9 9 10 10
3 3 5 6 7 7 7 10 10 11 12
3 3 5 6 7 7 8 8 9 9 9
3 3 5 6 7 7 8 9 9 12 12
3 3 5 6 7 7 9 9 12 12
3 3 5 6 7 7 9 10 10 10 13
3 3 5 6 7 7 9 10 11 12 12
3 3 5 6 7 8 8 11 12 12
3 3 5 6 7 8 8 10 10 10 13
3 3 5 6 7 8 8 10 11 12 12
3 3 5 6 7 8 9 9 11 11 13
3 3 5 6 7 8 9 10 11 11 11
3 4 4 4 5 6 6 6 7 7 7
3 4 4 4 5 6 6 7 7 8 8
3 4 4 4 5 6 7 7 7 9 9
3 4 4 4 5 6 7 7 8 9 11
3 4 4 4 5 6 7 7 9 11 11
3 4 4 4 5 6 7 7 11 12 12
3 4 4 4 5 6 7 9 10 12 12
3 4 4 4 5 6 7 10 11 11 13
3 4 4 4 6 6 7 7 7 9 12
3 4 4 4 6 6 7 7 10 11 11
3 4 4 4 6 7 7 7 8 8 11
3 4 4 4 6 7 7 7 9 10 11
3 4 4 4 6 7 7 7 11 12 12
3 4 4 4 6 7 7 8 8 10 12
3 4 4 4 6 7 7 8 10 10 13
3 4 4 4 6 7 7 8 11 12 12
3 4 4 4 6 7 7 8 10 10 13
3 4 4 4 6 7 7 9 10 10 13
3 4 4 4 6 7 7 10 10 10 13
3 4 4 4 6 7 7 10 11 12 12
3 4 4 4 6 7 8 9 9 9 9
3 4 4 4 6 7 8 9 9 12 12
3 4 4 4 6 7 9 9 9 12 12
3 4 4 4 6 7 9 10 10 10 13
3 4 4 4 6 7 9 10 11 12 12
3 4 4 4 6 7 9 10 10 10 11
3 4 4 4 7 7 7 10 10 11 12
3 4 4 4 7 7 7 10 11 12 12
3 4 4 4 7 7 8 9 9 11 12
3 4 4 4 7 7 8 9 9 11 12
3 4 4 4 7 7 9 9 11 11 13
3 4 4 4 7 7 9 10 11 11 11

3 3 5 6 6 7 8 8 9 9 13
3 3 5 6 6 7 8 8 11 12 12
3 3 5 6 6 7 8 10 11 11 11
3 3 5 6 6 8 8 11 11 13
3 3 5 6 6 8 8 9 10 12 12
3 3 5 6 6 8 8 10 11 11 13
3 3 5 6 6 8 9 9 11 11 11
3 3 5 6 6 8 9 10 10 12 12
3 3 5 6 7 7 7 9 10 10 10
3 3 5 6 7 7 7 10 10 12 12
3 3 5 6 7 7 8 8 9 9 13
3 3 5 6 7 7 8 9 11 11 11
3 3 5 6 7 7 9 9 10 11 11
3 3 5 6 7 7 9 10 10 11 12
3 3 5 6 7 8 8 8 9 10 11
3 3 5 6 7 8 8 9 10 10
3 3 5 6 7 8 8 10 10 11 12
3 3 5 6 7 8 9 9 11 12
3 3 5 6 7 8 9 9 11 11 13
3 4 4 4 5 6 6 7 7 7 8
3 4 4 4 5 6 6 7 7 10 12
3 4 4 4 5 6 7 7 7 9 12
3 4 4 4 5 6 7 7 9 9 9
3 4 4 4 5 6 7 7 9 12 12
3 4 4 4 5 6 7 8 9 10 10
3 4 4 4 5 6 7 10 10 10 13
3 4 4 4 5 6 7 10 11 12 12
3 4 4 4 6 6 7 7 7 11 12
3 4 4 4 6 6 7 7 11 11 11
3 4 4 4 6 7 7 7 8 8 13
3 4 4 4 6 7 7 7 10 11 11
3 4 4 4 6 7 7 8 8 8 10
3 4 4 4 6 7 7 8 8 11 12
3 4 4 4 6 7 7 8 10 11 11
3 4 4 4 6 7 7 9 10 10 10
3 4 4 4 6 7 7 9 10 10 12
3 4 4 4 6 7 8 8 9 9 9
3 4 4 4 6 7 8 9 9 12 12
3 4 4 4 6 7 9 10 10 10 13
3 4 4 4 6 7 9 10 11 12 12
3 4 4 4 7 7 7 9 10 10 11
3 4 4 4 7 7 7 10 10 11 12
3 4 4 4 7 7 8 8 9 9 9
3 4 4 4 7 7 8 9 9 12 12
3 4 4 4 7 7 9 9 9 12 12
3 4 4 4 7 7 9 9 11 12 12
3 4 4 4 7 7 9 10 11 11 13

3 3 5 6 6 7 8 8 9 12 12
3 3 5 6 6 7 8 9 9 10 10
3 3 5 6 6 7 8 10 11 12 12
3 3 5 6 6 8 8 8 11 12 12
3 3 5 6 6 8 8 10 10 10 13
3 3 5 6 6 8 8 10 11 12 12
3 3 5 6 6 8 9 9 11 11 13
3 3 5 6 6 8 9 10 11 11 11
3 3 5 6 6 8 9 10 10 13
3 3 5 6 7 7 7 9 10 10 13
3 3 5 6 7 7 7 10 11 11 11
3 3 5 6 7 7 8 9 9 9 13
3 3 5 6 7 7 8 9 11 11 13
3 3 5 6 7 7 9 9 11 11 11
3 3 5 6 7 7 9 10 10 12 12
3 3 5 6 7 8 8 8 9 10 11 11
3 3 5 6 7 8 8 9 10 10 10
3 3 5 6 7 8 8 10 10 12 12
3 3 5 6 7 8 9 9 12 12
3 3 5 6 7 8 9 10 10 13
3 3 5 6 7 8 9 10 11 12 12
3 4 4 4 5 6 6 7 7 7 10
3 4 4 4 5 6 6 7 7 11 12
3 4 4 4 5 6 7 7 7 11 12
3 4 4 4 5 6 7 7 9 9 11
3 4 4 4 5 6 7 7 10 11 11
3 4 4 4 5 6 7 10 10 10 13
3 4 4 4 5 6 7 10 11 12 12
3 4 4 4 6 6 7 7 7 12 12
3 4 4 4 6 6 7 7 11 11 13
3 4 4 4 6 7 7 7 8 10 11
3 4 4 4 6 7 7 7 10 11 13
3 4 4 4 6 7 7 8 8 13
3 4 4 4 6 7 7 8 10 12 12
3 4 4 4 6 7 7 8 10 12 12
3 4 4 4 6 7 7 9 10 10 10
3 4 4 4 6 7 7 10 10 12 12
3 4 4 4 6 7 8 8 9 9 13
3 4 4 4 6 7 8 9 11 11 11
3 4 4 4 6 7 9 9 10 11 11
3 4 4 4 6 7 9 10 10 11 12
3 4 4 4 7 7 7 8 9 10 10
3 4 4 4 7 7 7 9 10 10 13
3 4 4 4 7 7 8 8 9 9 13
3 4 4 4 7 7 8 9 11 11 11
3 4 4 4 7 7 9 9 12 13
3 4 4 4 7 7 9 10 10 10 13
3 4 4 4 7 7 9 10 11 12 12

3 4 4 4 7 8 8 8 9 10 11
3 4 4 4 7 8 8 8 11 12 12
3 4 4 4 7 8 8 10 10 10 13
3 4 4 4 7 8 8 10 11 12 12
3 4 4 4 7 8 9 9 11 11 11
3 4 4 4 7 8 9 10 10 12 12
3 4 4 5 5 6 6 6 11 12
3 4 4 5 5 6 6 6 8 8 8
3 4 4 5 5 6 6 6 8 11 13
3 4 4 5 5 6 6 6 11 11 11
3 4 4 5 5 6 6 8 8 10 10
3 4 4 5 5 6 6 9 9 10 10
3 4 4 5 5 6 6 10 10 10 13
3 4 4 5 5 6 6 10 11 12 12
3 4 4 5 6 6 6 7 7 10 12
3 4 4 5 6 6 6 7 9 10 10
3 4 4 5 6 6 6 7 10 12 12
3 4 4 5 6 6 6 8 9 10 13
3 4 4 5 6 6 6 9 10 11 12
3 4 4 5 6 6 6 10 11 11 11
3 4 4 5 6 6 7 7 9 12
3 4 4 5 6 6 7 7 8 9 11
3 4 4 5 6 6 7 7 9 11 11
3 4 4 5 6 6 7 7 11 12 12
3 4 4 5 6 6 7 8 9 10 10
3 4 4 5 6 6 7 9 9 11 12
3 4 4 5 6 6 7 9 11 11 11
3 4 4 5 6 6 7 10 10 12 12
3 4 4 5 6 6 8 8 9 9 13
3 4 4 5 6 6 8 9 9 11 12
3 4 4 5 6 6 8 9 11 12 12
3 4 4 5 6 6 9 9 10 11 13
3 4 4 5 6 6 9 9 10 11 13
3 4 4 5 6 6 9 10 10 11 12
3 4 4 5 6 6 7 7 8 8 8
3 4 4 5 6 7 7 8 8 12 12
3 4 4 5 6 7 8 8 9 9 9
3 4 4 5 6 7 8 8 11 11 11
3 4 4 5 6 7 8 10 10 11 12
3 4 4 5 6 8 8 8 9 10 11
3 4 4 5 6 8 8 8 11 12 12
3 4 4 5 6 8 8 9 10 11 12
3 4 4 5 6 8 8 10 11 11 11
3 4 4 5 6 8 8 9 9 12 13
3 4 4 5 6 8 9 9 11 12 12
3 4 4 5 6 8 9 10 11 11 13
3 4 4 5 6 8 9 10 10 12 12
3 4 4 5 6 8 9 10 11 12 12
3 4 4 6 6 6 7 8 9 9 13
3 4 4 6 6 6 8 9 9 10
3 4 4 6 6 6 8 9 9 12 12
3 4 4 6 6 6 8 9 11 12 12

3 4 4 4 7 8 8 8 10 11 11
3 4 4 4 7 8 8 9 9 10 10
3 4 4 4 7 8 8 10 10 11 12
3 4 4 4 7 8 9 9 9 11 12
3 4 4 4 7 8 9 9 11 11 13
3 4 4 4 7 8 9 10 11 11 11
3 4 4 5 5 6 6 6 12 12
3 4 4 5 5 6 6 6 8 8 11
3 4 4 5 5 6 6 6 8 12 12
3 4 4 5 5 6 6 6 11 11 13
3 4 4 5 5 6 6 8 9 10 10
3 4 4 5 5 6 6 9 10 10 10
3 4 4 5 5 6 6 10 10 11 12
3 4 4 5 6 6 6 7 7 7 10
3 4 4 5 6 6 6 7 7 10 13
3 4 4 5 6 6 6 7 9 10 13
3 4 4 5 6 6 6 7 11 11 11
3 4 4 5 6 6 6 9 9 10 10
3 4 4 5 6 6 6 9 10 11 12
3 4 4 5 6 6 6 10 11 11 13
3 4 4 5 6 6 7 7 7 11 12
3 4 4 5 6 6 7 7 9 9 9
3 4 4 5 6 6 7 7 9 12 12
3 4 4 5 6 6 7 8 8 9 9
3 4 4 5 6 6 7 8 9 11 12
3 4 4 5 6 6 7 9 9 12 12
3 4 4 5 6 6 7 9 11 11 13
3 4 4 5 6 6 7 10 11 11 11
3 4 4 5 6 6 8 9 9 9 10
3 4 4 5 6 6 8 9 9 12 12
3 4 4 5 6 6 8 9 9 11 12
3 4 4 5 6 6 9 9 9 11 12
3 4 4 5 6 6 9 9 11 11 11
3 4 4 5 6 6 9 10 10 12 12
3 4 4 5 6 6 7 7 8 8 8 10
3 4 4 5 6 7 7 8 8 8 9 9
3 4 4 5 6 7 8 8 8 9 9 11
3 4 4 5 6 7 8 8 11 11 13
3 4 4 5 6 7 8 10 10 12 12
3 4 4 5 6 8 8 8 10 11 11
3 4 4 5 6 8 8 8 9 9 10 10
3 4 4 5 6 8 8 8 9 10 12 12
3 4 4 5 6 8 8 10 11 11 13
3 4 4 5 6 8 8 9 10 11 11
3 4 4 5 6 8 9 10 10 13
3 4 4 5 6 8 9 10 11 12 12
3 4 4 6 6 6 7 8 8 9 9
3 4 4 6 6 6 7 8 9 11 12
3 4 4 6 6 6 8 9 9 13
3 4 4 6 6 6 8 9 9 12 13
3 4 4 6 6 6 8 9 9 11 12

3 4 4 4 7 8 8 8 10 11 13
3 4 4 4 7 8 8 9 10 10 10
3 4 4 4 7 8 8 10 10 12 12
3 4 4 4 7 8 9 9 9 12 12
3 4 4 4 7 8 9 9 11 12 12
3 4 4 4 7 8 9 10 11 11 13
3 4 4 5 5 6 6 6 12 13
3 4 4 5 5 6 6 6 8 8 13
3 4 4 5 5 6 6 6 9 10 11
3 4 4 5 5 6 6 6 11 12 12
3 4 4 5 5 6 6 8 10 10 10
3 4 4 5 5 6 6 9 10 10 11
3 4 4 5 5 6 6 10 10 12 12
3 4 4 5 6 6 6 7 7 7 13
3 4 4 5 6 6 6 7 7 11 12
3 4 4 5 6 6 6 7 10 10 10
3 4 4 5 6 6 6 7 11 11 13
3 4 4 5 6 6 6 9 10 10 10
3 4 4 5 6 6 6 10 10 10 13
3 4 4 5 6 6 6 10 11 12 12
3 4 4 5 6 6 7 7 7 12 12
3 4 4 5 6 6 7 7 9 9 11
3 4 4 5 6 6 7 7 10 11 11
3 4 4 5 6 6 7 8 9 9 9
3 4 4 5 6 6 7 8 9 12 12
3 4 4 5 6 6 7 9 10 10 10
3 4 4 5 6 6 7 9 11 12 12
3 4 4 5 6 6 7 10 11 11 13
3 4 4 5 6 6 8 9 9 9 13
3 4 4 5 6 6 8 9 9 10 10
3 4 4 5 6 6 8 9 10 11 11
3 4 4 5 6 6 8 9 9 12 12
3 4 4 5 6 6 9 9 9 12 12
3 4 4 5 6 6 9 9 11 11 13
3 4 4 5 6 6 9 10 11 11 11
3 4 4 5 6 6 9 10 11 11 13
3 4 4 5 6 6 9 10 11 12 12
3 4 4 5 6 6 7 7 8 8 8 13
3 4 4 5 6 7 8 8 8 11 12
3 4 4 5 6 7 8 8 9 9 13
3 4 4 5 6 7 8 8 11 12 12
3 4 4 5 6 7 8 8 11 12 12
3 4 4 5 6 7 8 10 11 11 11
3 4 4 5 6 8 8 8 10 11 13
3 4 4 5 6 8 8 9 10 10 10
3 4 4 5 6 8 8 10 10 10 13
3 4 4 5 6 8 8 10 11 12 12
3 4 4 5 6 8 8 9 10 11 13
3 4 4 5 6 8 9 10 10 11 12
3 4 4 6 6 6 7 8 8 9 9
3 4 4 6 6 6 7 8 9 12 12
3 4 4 6 6 6 8 8 9 9 9
3 4 4 6 6 6 8 9 9 10 10
3 4 4 6 6 6 8 9 10 11 11
3 4 4 6 6 6 8 9 9 12 12
3 4 4 6 6 6 9 9 9 12 12

3 4 4 4 7 8 8 8 11 11 11
3 4 4 4 7 8 8 9 10 10 13
3 4 4 4 7 8 8 10 11 11 11
3 4 4 4 7 8 9 9 9 12 13
3 4 4 4 7 8 9 10 10 10 13
3 4 4 4 7 8 9 10 11 12 12
3 4 4 5 5 6 6 6 7 8 8
3 4 4 5 5 6 6 6 8 10 11
3 4 4 5 5 6 6 6 10 11 11
3 4 4 5 5 6 6 6 7 8 9 10
3 4 4 5 5 6 6 8 10 10 13
3 4 4 5 5 6 6 9 10 10 13
3 4 4 5 5 6 6 10 10 11 11
3 4 4 5 6 6 6 7 7 9 10
3 4 4 5 6 6 6 7 7 12 12
3 4 4 5 6 6 6 7 10 10 13
3 4 4 5 6 6 6 7 11 12 12
3 4 4 5 6 6 6 9 10 10 11
3 4 4 5 6 6 6 10 10 11 12
3 4 4 5 6 6 7 7 7 8 9
3 4 4 5 6 6 7 7 7 12 13
3 4 4 5 6 6 7 7 9 9 13
3 4 4 5 6 6 7 7 11 11 11
3 4 4 5 6 6 7 8 9 9 11
3 4 4 5 6 6 7 8 9 12 12
3 4 4 5 6 6 7 9 10 10 13
3 4 4 5 6 6 7 10 10 10 13
3 4 4 5 6 6 7 10 11 12 12
3 4 4 5 6 6 8 9 9 10 10
3 4 4 5 6 6 8 9 11 11 11
3 4 4 5 6 6 8 9 9 12 13
3 4 4 5 6 6 9 9 11 12 12
3 4 4 5 6 6 9 10 11 11 13
3 4 4 5 6 6 9 10 11 11 13
3 4 4 5 6 6 9 10 11 12 12
3 4 4 5 6 6 7 7 8 8 10 12
3 4 4 5 6 7 8 8 8 12 12
3 4 4 5 6 7 8 8 9 12 12
3 4 4 5 6 7 8 8 9 10 10
3 4 4 5 6 7 8 8 9 11 11 13
3 4 4 5 6 7 8 8 10 11 11
3 4 4 5 6 7 8 10 11 11 13
3 4 4 5 6 8 8 8 11 11 11
3 4 4 5 6 8 8 9 10 10 11
3 4 4 5 6 8 8 10 10 11 12
3 4 4 5 6 8 8 9 9 11 12
3 4 4 5 6 8 9 9 11 11 11
3 4 4 5 6 8 9 10 10 12 12
3 4 4 6 6 6 7 8 9 9 9
3 4 4 6 6 6 8 8 9 9 9
3 4 4 6 6 6 8 8 9 10 12
3 4 4 6 6 6 8 9 11 11 11
3 4 4 6 6 6 9 9 9 12 13

3 4 4 4 7 8 8 8 11 11 13
3 4 4 4 7 8 8 9 10 12 12
3 4 4 4 7 8 8 10 11 11 13
3 4 4 4 7 8 9 9 10 11 11
3 4 4 4 7 8 9 10 10 11 12
3 4 4 5 5 6 6 6 9 12
3 4 4 5 5 6 6 6 7 8 11
3 4 4 5 5 6 6 6 8 11 11
3 4 4 5 5 6 6 6 10 11 13
3 4 4 5 5 6 6 6 8 8 9 10
3 4 4 5 5 6 6 8 10 12 12
3 4 4 5 5 6 6 9 10 12 12
3 4 4 5 5 6 6 10 11 11 13
3 4 4 5 6 6 6 7 7 10 10
3 4 4 5 6 6 6 7 8 9 10
3 4 4 5 6 6 6 7 10 11 11
3 4 4 5 6 6 6 8 9 10 10
3 4 4 5 6 6 6 10 10 12 12
3 4 4 5 6 6 7 7 7 9 9
3 4 4 5 6 6 7 7 8 9 9
3 4 4 5 6 6 7 7 9 10 11
3 4 4 5 6 6 7 7 11 11 13
3 4 4 5 6 6 7 8 9 9 13
3 4 4 5 6 6 7 9 9 10 10
3 4 4 5 6 6 7 9 10 12 12
3 4 4 5 6 6 7 10 10 11 12
3 4 4 5 6 6 8 8 9 9 9
3 4 4 5 6 6 8 9 9 10 12
3 4 4 5 6 6 8 9 11 11 13
3 4 4 5 6 6 9 9 10 11 11
3 4 4 5 6 6 9 10 10 13
3 4 4 5 6 6 9 10 11 12 12
3 4 4 5 6 6 7 7 8 8 11 12
3 4 4 5 6 7 8 8 8 12 13
3 4 4 5 6 7 8 8 10 11 11
3 4 4 5 6 7 8 10 10 10 13
3 4 4 5 6 7 8 10 11 12 12
3 4 4 5 6 8 8 8 11 11 13
3 4 4 5 6 8 8 8 9 10 13
3 4 4 5 6 8 8 9 10 13
3 4 4 5 6 8 9 9 11 11 13
3 4 4 5 6 8 9 10 11 11 11
3 4 4 6 6 6 7 8 9 9 11
3 4 4 6 6 6 8 8 9 9 13
3 4 4 6 6 6 8 9 9 11 12
3 4 4 6 6 6 8 9 11 13
3 4 4 6 6 6 9 9 10 11 11

3 4 5 6 6 7 8 8 11 12 12	3 4 5 6 6 7 8 8 11 12 13	3 4 5 6 6 7 8 9 9 10 10	3 4 5 6 6 7 8 9 10 10 10	3 4 5 6 6 7 8 9 10 10 11
3 4 5 6 6 7 8 9 10 10 13	3 4 5 6 6 7 8 9 10 11 12	3 4 5 6 6 7 8 9 10 12 12	3 4 5 6 6 7 8 9 10 12 13	3 4 5 6 6 7 8 10 10 10 13
3 4 5 6 6 7 8 10 10 11 12	3 4 5 6 6 7 8 10 10 12 12	3 4 5 6 6 7 8 10 10 12 13	3 4 5 6 6 7 8 10 11 11 11	3 4 5 6 6 7 8 10 11 11 13
3 4 5 6 6 7 8 10 11 12 12	3 4 5 6 6 8 8 8 9 10 11	3 4 5 6 6 8 8 8 10 11 11	3 4 5 6 6 8 8 10 11 13	3 4 5 6 6 8 8 11 11 11
3 4 5 6 6 8 8 8 11 11 13	3 4 5 6 6 8 8 8 11 12 12	3 4 5 6 6 8 8 8 11 12 13	3 4 5 6 6 8 8 9 9 10 10	3 4 5 6 6 8 8 9 10 10 10
3 4 5 6 6 8 8 9 10 10 11	3 4 5 6 6 8 8 9 10 10 13	3 4 5 6 6 8 8 9 10 11 12	3 4 5 6 6 8 8 9 10 12 12	3 4 5 6 6 8 8 9 10 12 13
3 4 5 6 6 8 8 10 10 10 13	3 4 5 6 6 8 8 10 10 11 12	3 4 5 6 6 8 8 10 10 12 12	3 4 5 6 6 8 8 10 10 12 13	3 4 5 6 6 8 8 10 11 11 11
3 4 5 6 6 8 8 10 11 11 13	3 4 5 6 6 8 8 10 11 12 12	3 4 5 6 6 8 8 10 11 12 13	3 4 5 6 6 8 9 9 11 12	3 4 5 6 6 8 9 9 12 12
3 4 5 6 6 8 9 9 12 13	3 4 5 6 6 8 9 9 10 11 11	3 4 5 6 6 8 9 9 10 11 13	3 4 5 6 6 8 9 9 11 11 11	3 4 5 6 6 8 9 9 11 11 13
3 4 5 6 6 8 9 9 11 12 12	3 4 5 6 6 8 9 9 11 12 13	3 4 5 6 6 8 9 10 10 10 13	3 4 5 6 6 8 9 10 10 11 12	3 4 5 6 6 8 9 10 10 12 12
3 4 5 6 6 8 9 10 10 12 13	3 4 5 6 6 8 9 10 11 11 11	3 4 5 6 6 8 9 10 11 11 13	3 4 5 6 6 8 9 10 11 12 12	3 4 5 6 7 7 7 8 9 10 10
3 4 5 6 7 7 7 8 9 10 13	3 4 5 6 7 7 7 9 10 10 10	3 4 5 6 7 7 7 9 10 10 10	3 4 5 6 7 7 7 9 10 10 11	3 4 5 6 7 7 7 9 10 10 13
3 4 5 6 7 7 7 9 10 11 12	3 4 5 6 7 7 7 9 10 12 12	3 4 5 6 7 7 7 9 10 12 13	3 4 5 6 7 7 7 10 10 10 13	3 4 5 6 7 7 7 10 10 11 12
3 4 5 6 7 7 7 10 10 12 12	3 4 5 6 7 7 7 10 10 12 13	3 4 5 6 7 7 7 10 11 11 11	3 4 5 6 7 7 7 10 11 11 13	3 4 5 6 7 7 7 10 11 12 12
3 4 5 6 7 7 7 10 11 12 13	3 4 5 6 7 7 8 8 9 9 9	3 4 5 6 7 7 8 8 9 9 13	3 4 5 6 7 7 8 9 9 10	3 4 5 6 7 7 8 9 9 13
3 4 5 6 7 7 8 9 9 10 10	3 4 5 6 7 7 8 9 9 10 12	3 4 5 6 7 7 8 9 9 10 13	3 4 5 6 7 7 8 9 9 11 12	3 4 5 6 7 7 8 9 9 12 12
3 4 5 6 7 7 8 9 9 12 13	3 4 5 6 7 7 8 9 10 11 11	3 4 5 6 7 7 8 9 10 11 13	3 4 5 6 7 7 8 9 11 11 11	3 4 5 6 7 7 8 9 11 11 13
3 4 5 6 7 7 8 9 11 12 12	3 4 5 6 7 7 8 9 11 12 13	3 4 5 6 7 7 9 9 11 12	3 4 5 6 7 7 9 9 12 12	3 4 5 6 7 7 9 9 12 13
3 4 5 6 7 7 9 10 11 11	3 4 5 6 7 7 9 10 11 13	3 4 5 6 7 7 9 11 11 11	3 4 5 6 7 7 9 11 11 13	3 4 5 6 7 7 9 11 12 12
3 4 5 6 7 7 9 11 12 13	3 4 5 6 7 7 9 10 10 10 13	3 4 5 6 7 7 9 10 10 11 12	3 4 5 6 7 7 9 10 10 12 12	3 4 5 6 7 7 9 10 10 12 13
3 4 5 6 7 7 9 10 11 11 11	3 4 5 6 7 7 9 10 11 11 13	3 4 5 6 7 7 9 10 11 12 12	3 4 5 6 7 7 9 10 11 12 13	3 4 5 6 7 8 8 8 9 10 11
3 4 5 6 7 8 8 8 10 11 11	3 4 5 6 7 8 8 8 10 11 13	3 4 5 6 7 8 8 8 11 11 11	3 4 5 6 7 8 8 11 11 13	3 4 5 6 7 8 8 8 11 12 12
3 4 5 6 7 8 8 8 11 12 13	3 4 5 6 7 8 8 9 9 10 10	3 4 5 6 7 8 8 9 10 10 10	3 4 5 6 7 8 8 9 10 10 11	3 4 5 6 7 8 8 9 10 10 13
3 4 5 6 7 8 8 9 10 11 12	3 4 5 6 7 8 8 9 10 12 12	3 4 5 6 7 8 8 9 10 12 13	3 4 5 6 7 8 8 10 10 10 13	3 4 5 6 7 8 8 10 10 11 12
3 4 5 6 7 8 8 10 10 12 12	3 4 5 6 7 8 8 10 10 12 13	3 4 5 6 7 8 8 10 11 11 11	3 4 5 6 7 8 8 10 11 11 13	3 4 5 6 7 8 8 10 11 12 12
3 4 5 6 7 8 8 10 11 12 13	3 4 5 6 7 8 9 9 11 12	3 4 5 6 7 8 9 9 12 12	3 4 5 6 7 8 9 9 12 13	3 4 5 6 7 8 9 9 10 11 11
3 4 5 6 7 8 9 9 10 11 13	3 4 5 6 7 8 9 9 11 11 11	3 4 5 6 7 8 9 9 11 11 13	3 4 5 6 7 8 9 9 11 12 12	3 4 5 6 7 8 9 9 11 12 13
3 4 5 6 7 8 9 10 10 10 13	3 4 5 6 7 8 9 10 10 11 12	3 4 5 6 7 8 9 10 10 12 12	3 4 5 6 7 8 9 10 10 12 13	3 4 5 6 7 8 9 10 11 11 11
3 4 5 6 7 8 9 10 11 11 13	3 4 5 6 7 8 9 10 11 12 12	3 4 5 6 7 8 9 10 11 12 13		

Figure 6. LDF codes of all 1858 acyclic alkane isomers. Note that the first two entries (1 2) of all codes are identical, and have been omitted.

RESULTS AND DISCUSSION

The LDF codes of all 1858 acyclic alkane isomers are listed in Figure 6. Since each LDF code starts with entries 1 and 2, these have been omitted from the LDF codes in Figure 5. Therefore the code representing the structure in Figure 4 appears in Figure 6 as a string of the following numbers: 3,3,5,5,5,7,8,8,10,11,11. (In Figure 6 each comma is replaced by a blank.) The last code (1,2,)3,4,5,6,7,8,9,10,11,12,13 represents the normal chain derivative, n-quatuordecane.

The present rules are sufficient to delete all redundant structures from the set of Morgan trees. The selection procedure can be executed in quadratic time. No comparisons of different structures are necessary. Although computer time might be reduced by devising still more efficient algorithms, this would hardly affect the overall performance. In order to support this assumption let us suppose that someone created an algorithm, by which the structures of isomers are generated directly (i.e. without any selection procedure) and the generation procedure is accomplished in zero time. Even in this (impossible) case the time needed for printing of the results would increase exponentially, since the number of isomers increases exponentially. It must be mentioned that the method can be used to generate codes of all radicals or singly substituted alkane derivatives, too.

References

- [1] H. R. Henze, C. M. Blair, *The Number of Isomeric Hydrocarbons of the Methane Series*, J. Amer. Chem. Soc. **53** (1931) 3077-3085.
- [2] J. V. Knop, W. R. Müller, Z. Jeričević, N. Trinajstić, *Computer Enumeration and Generation of Trees and Rooted Trees*, J. Chem. Inf. Comput. Sci. **21** (1981) 91-99.
- [3] M. L. Contreras, J. Alvarez, M. Riveros, G. Arias, R. Rozas, *Exhaustive Generation of Organic Isomers. 6. Stereoisomers Having Isolated and Spiro Cycles and New Extended N-Tuples*, J. Chem. Inf. Comput. Sci. **41** (2001) 964-977.
- [4] I. Lukovits, *Isomer Generation: Syntactic Rules for Detection of Isomorphism*, J. Chem. Inf. Comput. Sci. **39** (1999) 563-568.
- [5] I. Lukovits, *Isomer Generation: Semantic Rules for Detection of Isomorphism*, J. Chem. Inf. Comput. Sci. **40** (2000) 361-366.
- [6] I. Lukovits, I. Gutman: *On Morgan-trees*, Croat. Chem. Acta, **75** (2002) 563-576.
- [7] I. Lukovits, in *Chemical Graph Theory* (Proceedings of the Harry Wiener Memorial Conference, Athens, Ga, D. H. Rouvray and R. B. King, eds) Horwood, Chichester, 2001, pp. 327-337.