

Constructive Enumeration of Chiral Isomers of Alkanes*

István Lukovits

Chemical Research Center, H-1525 Budapest, P. O. Box 17, Hungary
(E-mail: lukovits@chemres.hu)

RECEIVED APRIL 2, 2003; REVISED AUGUST 25, 2003; ACCEPTED SEPTEMBER 18, 2003

Key words
chiral isomers of alkanes
generation of structural formulas
codes of Morgan trees
canonical numbering

The problem of exhaustive and non-redundant generation of formulas of chiral isomers of alkanes was considered. The procedure starts with the generation of optimal codes of all *N*-alkane structures (*N* denotes the number of carbon atoms.) and then the algorithm determines whether a given structure contains at least a single carbon atom, with four different substituents. In order to accomplish these tasks, two simple rules have been used to delete redundant structures and determine whether two (alkyl) branches are identical or not.

INTRODUCTION

The problem of enumerating structural isomers has fascinated mathematicians and chemists since the end of the nineteenth century. The most basic problems have already been solved by the end of the thirties of the last century. Nevertheless, since then a great number of researchers have addressed this problem and the latest developments have been reviewed recently.^{1,2,3} The formal enumeration of isomers is most often accomplished by using generating functions.^{1,3} The problem of the exhaustive and non-redundant generation of structural formulas of isomers⁴ has also attracted a great number of chemists, and one of the first reviews on this topic was written by Trinajstić *et al.* in 1985.⁵ Their method used to generate formulas of alkane isomers was based on the *N*-tuple representation (NTR).⁶ Since then the method based on NTR has been extended to include cycle containing structures, stereoisomers and structures containing heteroatoms.⁷ A computer program capable of generating the NTRs of any acyclic structure has also been created.⁸

Besides the NTR method, other approaches⁹ of generating formulas have also been proposed. Canonical la-

beling and coding of alkane isomers based on local vertex invariants and topological indices was proposed by Balaban *et al.*¹⁰ Because of the degeneracy in the topological indices, special care has to be devised to eliminate redundant structures in this approach. Davidson generated an alkane series dictionary, ordered by side chain complexity.¹¹ The alkane isomers were coded by using a special alphabet representing the simplest alkyl groups. Structure generation from fragments was proposed by Bohanec.¹² In this approach, however, duplicates could not be avoided, making comparisons of final structures necessary. Construction and enumeration of alkane isomers was also accomplished by Jackson and Bieber¹³ through utilizing the degree distribution domains and sub-domains of constitutional isomers. The concept of stereoisomerism was extended to knots and the topic was reviewed by Dobrowolski recently.¹⁴

Our approach,¹⁵ proposed earlier, is based on the generation of codes of Morgan-trees¹⁶ (MTs) and the concept of canonical numbering.³ Each code represents a (labeled) structure. The generated codes are checked by using simple rules to delete non-canonically labeled

* Dedicated to Professor Nenad Trinajstić on the occasion of his 65th birthday.

structures.¹⁵ Note that two rules (to be discussed below) are sufficient to determine whether the tree under consideration is numbered canonically, or not: there is no need to compare different structures. The number¹⁷ of MTs is just a small fraction of all possible physical trees.¹⁸ (Each vertex k of a physical tree – except vertex 1 – has just single neighbor labeled with an ordinal m being less than k .⁵) It was shown that the formula used to determine the number of Morgan-trees can also be derived by using the concept of Dyck-paths.¹⁹

In this paper the concept of canonical labeling will be extended to alkyl groups and the method will be used to generate formulas of optical isomers of (acyclic) alkanes. The problem is to determine whether the four substituents attached to carbon atom, m , are all different or not. Those alkyl-groups the vertices of which are labeled with numbers greater than m , can be evaluated relatively easily, the real problem is the branch where the numbering starts. The algorithm proposed in this work generates a complete and non-redundant set of optimal codes representing the structures of alkanes¹⁵ from which all structures containing at least one chiral carbon are singled out. The structure of this paper is the following: the next section introduces the »lowest degrees first« (LDF) concept and the canonical numbering based on this principle. In the third section we discussed and proved two rules used to determine whether a given labeling is canonical. These concepts have been used to generate the exhaustive and non-redundant set of structures representing both structural as well as stereoisomers. In the fourth section the selection of the stereoisomers out of this set was described and the results were discussed.

CANONICAL NUMBERING

Figure 1 illustrates four different methods of labeling used in this work. Structure a (Figure 1) represents an unconstrained labeling, with no naming rules to be taken into account. Structure b (Figure 1) is an example of a »physical tree«: the technical term was coined by Trinajstić *et al.*,⁵ and means that numbers $2, \dots, N$ (N denotes the number of carbons) are assigned consecutively to vertices. At this point label 1 may be assigned to any vertex. Each consecutive new label n ($1 < n \leq N$) can be attached only to a vertex which has an already labeled neighbor. With this restriction the adjacency matrix of the structure becomes rather simple: each column of the right-hand triangle will contain just a single non-zero entry.¹⁷ (Since matrix A is symmetric, the same rule applies to the rows of the left-hand triangular part, too.) This fact in turn allows us to replace the adjacency matrix by the »condensed representation of matrix A « (CAM), which is a vector composed of $N-1$ entries. The abbreviation CAM refers to the denotation compressed

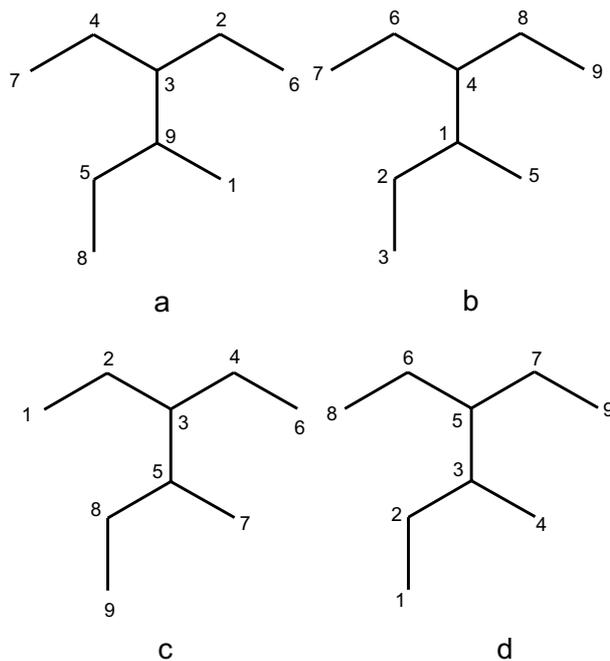


Figure 1. Four different numberings of the hydrogen suppressed graph of 3-methyl-4-ethyl-hexane: a, unconstrained labeling; b, physical tree; c, Morgan-tree; d, canonically labeled (LDF) tree.

adjacency matrix.¹⁷ Denoting the CAM of matrix A by vector C , we can transform matrix A of any physical tree into vector C by using the following rule:

$$\text{If } A_{ij} = 1, \text{ then } C_{j-1} = i \quad (j = 2, 3, \dots, N) \quad (1)$$

Note that $C_1 = 1$, in all cases. Examples have been given in earlier papers.^{15,17,18} Of course matrix A can readily be reconstructed from vector C :

$$\text{If } C_i = k, \text{ then } A_{k,i+1} = A_{i+1,k} = 1 \quad (2)$$

Example: if $C_1 = 1$, then $A_{1,2} = A_{2,1} = 1$, in accordance with the fact that in physical trees vertex 2 must be a neighbor of vertex 1.

A CAM can be determined directly by inspecting the numbering of a physical tree. As an example consider structure b in Figure 1. Inspect vertices 2, 3, 4, 5, 6, 7, 8, and 9 in turn, and observe the only first neighbor, which is labeled by a lower ordinal. This is vertex 1, vertex 2, vertex 1, vertex 1, vertex 4, vertex 6, vertex 4, and vertex 8, respectively, and therefore $C = (1, 2, 1, 1, 4, 6, 4, 8)$.

A Morgan-tree (MT) is a physical-tree which is numbered so, that $C_i \leq C_{i+1}$. It is clear that the number of MTs is (much) less than the number of physical trees.^{15,18} Structure c in Figure 1 is an example of a MT. In order to obtain a MT, the rule according to which each new label m should be attached to an already labeled vertex, must be supplemented by a second rule: label m must be attached to the vertex which is adjacent to a vertex labeled with the lowest ordinal. The CAM of

structure *c* (Figure 1) is $C' = (1,2,3,3,4,5,5,8)$, which can be obtained directly by inspecting structure *c*. Let's denote the valence of vertex *i* by d_i . Then the following simpler (than in the case of physical trees) algorithm can be used:

1. Set C_1 equal to 1.
2. If $d_1 = 1$, go to next step, otherwise write down 1 $d_1 - 1$ times.
3. If $d_2 = 1$, go to next step, otherwise write down 2 $d_2 - 1$ times.
4. Continue this procedure in the same manner until vertex $N - 1$ has been processed. Vertex N can not appear in C , since it must be an endpoint.

The reconstruction of structure *c* from its CAM, C' , can be accomplished in the same manner as in the case of physical-trees.

In what follows the method of lexicographic comparisons⁵ of vectors will be explained. Compare CAMs (or any other vector) C and C' . $C_1 = C'_1$ in all physical trees. If $C_2 = C'_2$, inspect C_3 and C'_3 . If $C_2 < C'_2$, then $C < C'$, etc. Formally, vectors composed of different number of entries can also be compared, if the last non-zero entry is followed by a number of zeros in the vector containing a fewer number of entries. Lexicographic comparisons will be utilized in the next section.

A »lowest degrees first« (LDF) tree is a MT (and therefore a physical-tree, too), in which a new label *n* is attached to a vertex, which has a first neighbor labeled with the lowest ordinal *k*, and which in addition has the lowest degree (valence) of all possible vertices attached to the same vertex *k*. Structure *d* (Figure 1) is an example of a LDF tree, and its CAM is equal to $C'' = (1,2,3,3,5,5,6,7)$. Note the lexicographic order of CAMs related to structures *b*, *c*, and *d* (Figure 1) C , C' , and C'' , respectively. The order is $C < C' < C''$.

In LDF trees the numbering must start at an endpoint. LDF trees are therefore canonically labeled trees as they possess the maximal CAM code of all MTs denoting the same structure (graph) *G*. An optimally coded tree is a tree which is numbered canonically with respect to a specified endpoint. An optimally numbered tree is therefore not necessarily a canonically numbered (*i.e.* LDF) tree, but an LDF tree is also an optimally numbered tree.

DETECTION OF NON-CANONICAL CODES

As it can be seen, different MTs may denote the same structure, and any tree can be labeled in a way that we obtain an MT. First all codes representing MTs of all structures (acyclic trees) composed of *N* vertices were generated in a systematic way, and then we deleted all codes representing non-canonically labeled structures. In fact even this procedure could be simplified, since many

MTs not representing LDFs were excluded from the generation procedure *a priori*. Details are given elsewhere.¹⁵

Systematic generation of codes of MTs may start with code $C = (1,1,1,1,\dots,1)$, which represents a star, the next code being equal to $C = (1,1,1,1,\dots,2)$, etc. The last code to be generated is $C = (1,2,3,\dots,N-1)$ representing an *N*-chain. In fact the procedure does not start with code $C = (1,1,1,1,\dots,1)$, but with $C = (1,2,2,2,\dots,2)$, since there is no LDF with a CAM less than $C = (1,2,2,2,\dots,2)$. There are still other rules, which can be used to exclude several codes greater than $C = (1,2,2,2,\dots,2)$ from the process without further processing them.¹⁵ Only those codes (structures) were generated in which the numbering starts at an endpoint.

Basically there are two rules, which are sufficient to determine whether a code (related to a MT) is canonical (*i.e.* it is an LDF code) or not. Both have been proved earlier,^{15,21} but the proof given in this work is much easier and simpler than earlier arguments. In order to prepare these arguments we have to introduce the concept of the valence code and the path code.

A valence code V is a vector containing *N* entries, and $V_i = d_i$, where d_i denotes the degree of vertex *i*. The valence code of structure *c* (Figure 1) is equal to $V = (1,2,3,2,3,1,1,2,1)$. It is clear that at least in the case of physical trees (therefore also for MTs and LDF trees), V can be obtained from C by, first, obtaining the adjacency matrix A from C , and then determining V by calculating the sum of entries of each row (or column) of matrix A . On the other hand we can obtain (at least in the case of MTs in which the numbering starts with an endpoint) vector C directly from vector V by using the following procedure.

1. Set C_1 equal to 1.
2. If $d_2 = 1$, go to next point, otherwise write down number 2 $d_2 - 1$ times.
3. If $d_3 = 1$, go to next point, otherwise write down number 3 $d_3 - 1$ times.
4. Continue until vertex $N - 1$ has been processed. Vertex N can not appear in C , since it must be an endpoint.

As an example let us consider the valence code of structure *c* (Figure 1) $V = (1,2,3,2,3,1,1,2,1)$, and convert it into vector C . By executing steps 1–4, we readily obtain $C = (1,2,3,3,4,5,5,8)$, which is identical with C' obtained by direct inspection of structure *c* (see previous chapter).

Of all MTs related to the same structure *G*, the LDF tree will have the lowest possible valence code by definition. Therefore because of the transformation algorithm delineated above, the corresponding CAM will be maximal, and this completes the proof of Theorem 1.

Theorem 1. – Of all possible MTs related to the same tree, *T*, the LDF (or canonically numbered) tree will have the minimal valence code and the maximal CAM.

As an example compare structures c and d (Figure 1). The respective valence codes are $V_c = (1,2,3,2,3,1,1,2,1)$, $V_d = (1,2,3,1,3,2,2,1,1)$, and $C_c = (1,2,3,3,4,5,5,8)$, $C_d = (1,2,3,3,5,5,6,7)$. In fact $V_c > V_d$ and because of this, $C_c < C_d$.

The valence code rule can be used to determine whether branches t_1 , t_2 , and t_3 attached to carbon atom, m , have been ordered correctly, or not.¹⁵ By »correct« order we mean that CAMs of t_1 , t_2 , and t_3 increase in that order. To obtain CAMs for branches t_i ($i = 1, 2, 3$) is a relatively easy procedure. Select the vertex with the lowest label in t_i and replace it by the new label 1. The select the next lowest label, and replace it by the new label 2, etc. This step can always be accomplished since the lowest labels in t_1 , t_2 , and t_3 will all be greater than m . The same »renumbering« procedure can not be applied for t_0 , because t_0 contains the endpoint at which the numbering starts and vertex j which is adjacent to m has a label being less than m . In the former cases, however, V_t (V_t denotes the valence code of any branch) can also be obtained for t_1 , t_2 , and t_3 . Note that valence codes can also be used to determine whether t_1 , t_2 , and t_3 are all different, which property will be used in selecting all codes (structures) that represent chiral alkanes.

In order to compare also t_0 with t_1 , t_2 , and t_3 , respectively, the technique based on the path code, P , was proposed.²⁰ P is a matrix containing $e-1$ rows (e denotes the number of endpoints) and d columns, where d denotes the length of the longest path of all paths starting at vertex 1. $P_{i,j}$ is equal to the valence of the j -th vertex placed on path i . If the particular paths are ordered lexicographically, we obtain the path code (i.e. matrix) P . As an example consider structure c (Figure 1). It can be seen by inspection that matrix P is equal to

1-2-3-2-1-0

1-2-3-3-1-0

1-2-3-3-2-1

The zero entries have been added in order to obtain rows of equal length. Dashes have been inserted to separate the entries. The path code of structure d (Figure 1) is equal to

1-2-3-1-0-0

1-2-3-3-2-1

1-2-3-3-2-1

Note that the rows of these matrices are already ordered lexicographically, meaning that if $m < n$, then $\text{row}_m \leq \text{row}_n$. Let's assume that P and P' are ordered lexicographically. If row_i ($1 \leq i \leq m-1$) in P is equal to row_i' in P' , and $\text{row}_m < \text{row}_m'$, then $P < P'$. The ordering is similar to the ordering of CAMs (section 2).

Next we want to show that if, and only if, $P < P'$ then $V < V'$. First, we show that if $V < V'$, then $P < P'$. Let's introduce instead of subscripts i and j a combined subscript, o ; where $o = (e-1)(j-1) + i$. If $i = 1$ and $j = 1$,

then $o = 1$, if $i = e-1$ and $j = 1$, then $o = e-1$, if $i = 5$ and $j = 3$, then $o = 2(e-1) + 5$, etc. Relation $V < V'$ means that there is a pair of entries such that $V_i < V_i'$ and $V_k = V_k'$ ($1 \leq k \leq i-1$). Both V_i and V_i' appear first in the same position m of P and P' , respectively, meaning that $P_o = P_o'$ for $1 \leq o < m$. Therefore, if $V < V'$, then $P < P'$.

Conversely, let's assume that $P < P'$, with $\text{row}_i < \text{row}_i'$ ($1 \leq i \leq e-1$), and $\text{row}_k = \text{row}_k'$ and $k < i$. Then $\text{row}_i(k) < \text{row}_i'(k)$, and $\text{row}_i(m) = \text{row}_i'(m)$, with $1 \leq m < k$, where $\text{row}_i(k)$ denotes the k -th entry in row i . But then $\text{row}_i(k) = V_n$ and $\text{row}_i'(k) = V_n'$, with $V_p = V_p'$ and $1 \leq p < n$. Therefore, if $P < P'$, then $V < V'$, which completes this proof.

Theorem 2. – If $P < P'$ then $V < V'$.

It can be seen, that already the first row of the path code of structure d is less than the corresponding row related to structure c, so in accordance with valence code, structure d has a greater CAM than structure c (Figure 1). The CAM of structure d is the LDF (i.e. maximal) code.

Since path codes can also be derived for substructures, i.e. branches, the question whether t_0 is identical with t_1 , t_2 , or t_3 can easily be decided, provided that all rows of the respective matrices P_1 , P_2 , and P_3 , are ordered lexicographically. Let's assume that vertex m' in t_0 is adjacent to vertex m (which is under consideration) in T. The (lexicographically ordered) path code P_0 related to side chain t_0 can easily be obtained with respect to m' as starting vertex (or vertex 1' in the new labeling) and so the question whether t_0 precedes t_1 , or whether t_0 is identical with t_1 , t_2 , and t_3 , respectively, can be settled.^{15,18}

APPLICATION: GENERATION OF ALL CHIRAL ISOMERS OF ALKANES

The computer program performing the generation procedure, starts the calculations by generating the LDF codes of all N -vertex alkane isomers,^{15,18,20} and then picks out all chiral derivatives. In order to determine whether the code represents a chiral structure, all vertices of degree 3 or 4 have to be inspected. Let's denote the carbon atom under consideration by index m . The next task is to determine whether the alkyl-groups t_0 , t_1 , t_2 , and (in case of $d_m = 4$) t_3 attached to m are all different, or not. Comparison of t_1 with t_2 or t_1 with t_3 , or t_2 with t_3 (if $d_m = 4$) are done by using Theorem 1, since if the tree is an LDF tree than its branches are also optimally numbered and their respective valence codes will be minimal. Therefore the valence codes of all branches have to be compared. Two valence codes are identical if all entries are identical. If the computer program detects a pair of identical branches, then the procedure continues with the inspection of the next vertex. If no vertex with three or four different branches are detected, then the code (and the underlying structure) will be discarded.

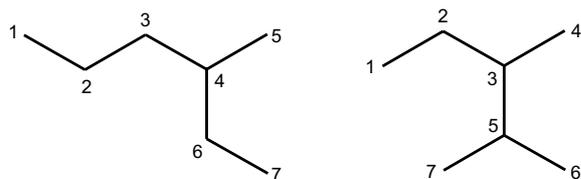


Figure 2. Heptanes possessing at least one chiral atom.

As was indicated in the previous section, comparison of branch t_0 with branches t_1 , t_2 , and t_3 (the latter only in case of a quaternary carbon atom) by using the valence codes does not work at all, and we have to compare the path codes.

Theorem 2 can be applied to determine whether t_0 is identical with any of the other substituents t_1 , t_2 , and t_3 . In chiral structures there must be at least one carbon atom, which four different substituents t_0 , t_1 , t_2 , and t_3 are attached to. The codes representing chiral structures were transformed into the corresponding formulas. Figure 2 lists the two optical isomers of heptane (with labeled vertices), while Figure 3 lists all five optical isomers of octane. The structures of chiral nonanes are shown in Figure 4. Each asymmetric carbon has been labeled by an asterisk. The number of isomers is therefore equal $(2 \times 12) + (4 \times 2) + (1 \times 3) = 35$, because structures 1, 2, 3, 4, 6, 8, 9, 10, 12, 13, 15 and 5, have to be taken into account two times, structures 5 and 11 have to be taken into account four times, and structure 7 represents three stereoisomers (*i.e.* RR, LL and RL). This result seems to be in contradiction with the result by Bytautas and Klein,¹ who claim that only 34 stereoisomers of nonanes exist. In fact there is no disagreement: Bytautas and Klein denote by »chiral« only optically active molecules, while the RL variant of structure 7 (Figure 4) is clearly optically inactive. The LDF codes of 40 chiral decane isomers are listed in Figure 5. For example the first code (1,2,2,4,4,6,6,8) represents 2,3,4-trimethyl-hexane, which has two asymmetric carbon atoms. The total number of structural isomers in the heptane–decane series of alkanes is 9, 18, 35, and 75, respectively, mean-

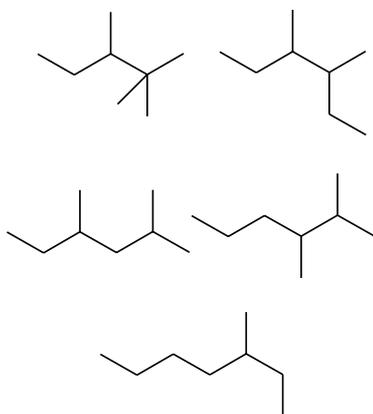


Figure 3. All octanes possessing at least one chiral atom.

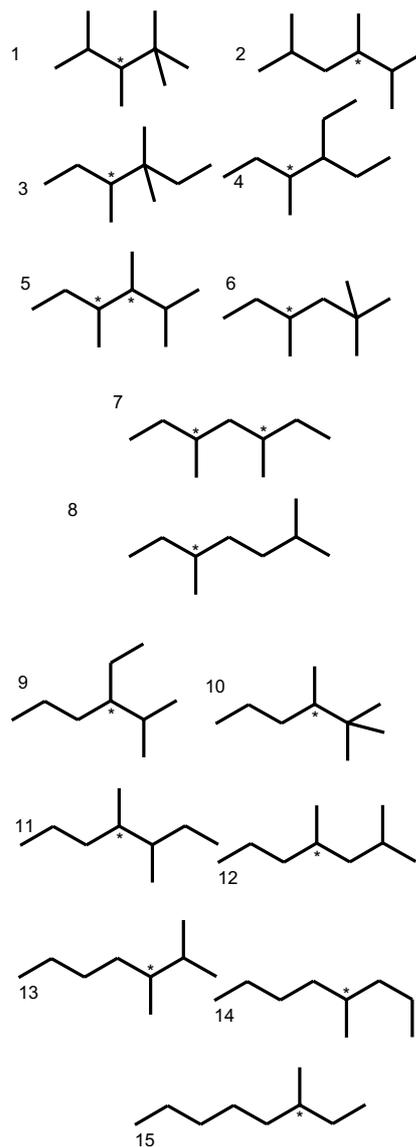


Figure 4. All nonanes possessing at least one chiral atom. The asymmetric carbons were labeled by an asterisk.

ing that in terms of N the ratio of chiral structures to the total number of structures increases monotonously. This result is in accordance with the conclusion by Bytautas and Klein.¹

The selection procedure (based on theorems 1 and 2) of codes representing the chiral structures can be performed in polynomial time. The same theorems and a set of additional, and simpler rules, which can be accomplished in linear time, were used to generate the list of codes representing all isomers of N -alkanes.¹⁵ It has been noted earlier, that only a small fraction of MTs had to be inspected, and a great portion of MTs was excluded from the screening procedure.¹⁸ On the other hand the number of N -alkane isomers increases exponentially in terms of N . This rule always holds if exhaustive generation of structures is attempted, even if an ideal »zero-time« procedure would be available, since in such cases the time

1 2 2 4 4 6 6 8 8
 1 2 2 4 4 6 7 7 7
 1 2 2 4 5 5 7 7 7
 1 2 2 4 5 6 6 8 8
 1 2 3 3 3 6 6 8 8
 1 2 3 3 4 4 5 5 5
 1 2 3 3 4 5 5 6 6
 1 2 3 3 4 5 5 8 8
 1 2 3 3 5 5 5 7 8
 1 2 3 3 5 5 5 8 8
 1 2 3 3 5 5 6 7 7
 1 2 3 3 5 5 7 7 7
 1 2 3 3 5 5 7 7 9
 1 2 3 3 5 5 7 8 8
 1 2 3 3 5 6 6 6 9
 1 2 3 3 5 6 6 7 8
 1 2 3 3 5 6 6 8 8
 1 2 3 3 5 6 7 7 7
 1 2 3 3 5 6 7 7 9
 1 2 3 3 5 6 7 8 8
 1 2 3 4 4 4 6 7 7
 1 2 3 4 4 4 7 7 9
 1 2 3 4 4 5 6 6 6
 1 2 3 4 4 5 6 6 9
 1 2 3 4 4 5 6 8 8
 1 2 3 4 4 6 6 6 9
 1 2 3 4 4 6 6 7 8
 1 2 3 4 4 6 6 8 8
 1 2 3 4 4 6 6 8 9
 1 2 3 4 4 6 7 7 7
 1 2 3 4 4 6 7 7 9
 1 2 3 4 4 6 7 8 8
 1 2 3 4 5 5 6 7 7
 1 2 3 4 5 5 6 7 9
 1 2 3 4 5 5 7 7 7
 1 2 3 4 5 5 7 7 9
 1 2 3 4 5 5 7 8 8
 1 2 3 4 5 6 6 8 8
 1 2 3 4 5 6 6 8 9
 1 2 3 4 5 6 7 7 9

Figure 5. LDF codes of all 40 decanes possessing at least one chiral atom.

needed for printing of results would still increase exponentially in terms of N . In cases where not all isomers have to be generated, linear time procedures are available. A particular example for such methods is the random generation of N -alkane isomers.^{17,21,22}

Acknowledgment. – This work was supported in part by Országos Tudományos Kutatási Alap (No. T043577).

REFERENCES

- (a) L. Bytautas and D. J. Klein, *J. Chem. Inf. Comput. Sci.* **39** (1999) 803–818; (b) L. Bytautas and D. J. Klein, *J. Chem. Inf. Comput. Sci.* **38** (1998) 1063–1078.
- D. J. Klein, D. Babić, and N. Trinajstić, in: A. Hinchliffe (Ed.), *Chemical Modelling: Applications and Theory*, Vol. 2. The Royal Society of Chemistry, London, 2002, pp. 56–95.
- R. C. Read, *Algorithms in Graph Theory*, in: A. T. Balaban (Ed.), *Chemical Applications of Graph Theory*, Academic Press, London, 1976, pp. 25–61.
- V. Kvasnička and J. Pospichal, *J. Chem. Inf. Comput. Sci.* **30** (1990) 99–105.
- J. V. Knop, W. R. Muller, K. Szymanski, S. Nikolić, and N. Trinajstić, *Computer Generation of Certain Classes of Molecules*, SKTH/Kemija u industriji, Zagreb, 1985.
- J. V. Knop, W. R. Müller, Z. Jeričević, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **21** (1981) 91–99.
- (a) M. L. Contreras, R. Rozas, and R. Valdivia, *J. Chem. Inf. Comput. Sci.* **34** (1994) 610–616; (b) M. L. Contreras, J. Alvarez, M. Riveros, G. Arias, and R. Rozas, *J. Chem. Inf. Comput. Sci.* **41** (2001) 964–977; (c) M. L. Contreras, G. M. Trevisol, J. Alvarez, G. Arias, and R. Rozas, *J. Chem. Inf. Comput. Sci.* **39** (1999) 475–482.
- P. Hansen, B. Jaumard, C. Lebatteux, and M. L. Zheng, *J. Chem. Inf. Comput. Sci.* **34** (1994) 782–790.
- C. Y. Hu, L. Xu, *Anal. Chim. Acta* **298** (1994) 75–85.
- T. S. Balaban, P. A. Filip, and O. Ivanciuc, *J. Math. Chem.* **11** (1992) 79–105.
- S. Davidson, *J. Chem. Inf. Comput. Sci.* **42** (2002) 147–156.
- S. Bohanec, *J. Chem. Inf. Comput. Sci.* **35** (1995) 494–503.
- M. D. Jackson and T. I. Bieber, *J. Chem. Inf. Comput. Sci.* **33** (1993) 701–708.
- J. Cz. Dobrowolski, *Croat. Chem. Acta* **76** (2003) 145–152.
- (a) I. Lukovits, *J. Chem. Inf. Comput. Sci.* **39** (1999) 563–568; (b) I. Lukovits, *J. Chem. Inf. Comput. Sci.* **40** (2000) 361–366.
- H. L. Morgan, *J. Chem. Doc.* **5** (1965) 107–113.
- I. Gutman, W. Linert, I. Lukovits, and Ž. Tomović, *J. Chem. Inf. Comput. Sci.* **40** (2000) 113–116.
- I. Lukovits and I. Gutman, *Croat. Chem. Acta* **75** (2002) 563–576.
- T. Došlić, *Croat. Chem. Acta* **75** (2002) 881–889.
- I. Lukovits, in: D. H. Rouvray and R. B. King (Eds.), *Topology in Chemistry*, Horwood Publishing Ltd., Chichester, 2002, pp. 333–343.
- L. Alonso, J. L. Remy, and R. Schott, *Algorithmica* **17** (1997) 162–182.
- M. D. Atkinson, *Comput. J.* **36** (1993) 593–594.

SAŽETAK

Konstruktivno prebrojavanje kiralnih izomera alkana

István Lukovits

Prikazan je postupak jednoznačnoga generiranja formula kiralnih izomera alkana. Postupak započinje generiranjem optimalnih kodova za alkane s određenim brojem ugljikovih atoma. Nakon toga se ispituje je li razmatrani alkan posjeduje bar jedan ugljikov atom, koji ima četiri supstituenta. Da bi se taj postupak mogao izvesti autor je dao dva jednostavna pravila pomoću kojih se uklanjaju suvišne strukture i određuje jesu li dvije alkilne grane identične ili ne.