shown, the last being similar to a plan of Dalton,[11] but no other morphine syntheses have used this concept. The cyclizations of the benzofurans in examples 1 and 7, however, seem to represent an especially efficient approach. The lower example in Figure 8 is the tricyclic terpene hirsutene, which yielded 9 and 18 conversions of three- and two-bond cyclizations, respectively. No examples of the use of multiple cyclizations exist in laboratory syntheses,[12] but the biosynthesis[13] does involve polycyclization of an 11-membered ring humulene precursor. The nearest such case produced here is example 15.

In summary, these two multiple construction programs offer new planning ideas for the efficient synthesis of a number of complex targets. Even when the detailed chemistry presented or implied here is not fully viable, variations incorporating the central theme may lead to successful syntheses that are often not likely to be conceived through common intuitive planning. Example 5 for hirsutene in Figure 8 illustrates this since the right-hand double bond has an added heteroatom placed on its lower carbon, whereas –OR on the upper carbon instead would provide a much better initiation of cyclization. To expand the potential of the SYNGEN program these two programs will now be added to it as modules offering a different aspect of the search for efficiency in synthesis design.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) Hendrickson, J. B.; Toczko, A. G. *J. Chem. Inf. Comput. Sci.* (preceding paper in this issue and references cited therein).
(2) This procedure creates four starting skeletons, which for a C_{20} target will average C_5 in size. Our experience with our starting materials catalog (some 6000 compounds) implies that the available functional group variety on the various skeletons falls off sharply above C_5. Therefore, it appears reasonable to expect of this procedure a fairly productive generation of routes from real catalog starting materials. A survey of published syntheses shows the average starting material to have only three skeletal carbons. Since this procedure will find routes from average C_5 compounds, it further serves to minimize the construction reactions required.
(3) Spencer, T. A.; Newton, M. D.; Baldwin, S. W. *J. Org. Chem.* **1964**, *29*, 787; improved by Ing-lung Shih in our laboratory.
(4) Desphande, M. N.; Jawdosiuk, M.; Kubiak, G.; Venkatachalam, M.; Weiss, V.; Cook, J. M. *J. Am. Chem. Soc.* **1985**, *107*, 4786, and references cited therein.
(5) Taub, D. In *Total Synthesis of Natural Products*; ApSimon, J., Ed.; Wiley: New York, 1973, Vol. 2, p 641; 1984, Vol. 6, p 1.
(6) Jung, M. E.; Halweg, K. M. *Tetrahedron Lett.* **1981**, 3929. Bal, S. A.; Helquist, P. *Ibid.* **1981**, 3939. Denmark, S. E.; Germanas, J. P. *Ibid.* **1984**, 1231.
(7) Stork, G., et al. *J. Am. Chem. Soc.* **1981**, *103*, 4948; **1982**, *104*, 3758.
(8) Hendrickson, J. B.; Braun-Keller, E. *J. Comput. Chem.* **1980**, *1*, 323.
(9) Posner, G. H. *Chem. Rev.* **1986**, *86*, 831.
(10) Johnson, W. S., et al. *J. Am. Chem. Soc.* **1976**, *98*, 1038, 1039; **1977**, *99*, 8341; **1982**, *104*, 3508. *Bioorg. Chem.* **1976**, *5*, 51.
(11) Dalton, D. R., private communication.
(12) Heathcock, C. H.; Graham, S. L.; Pirrung, M. C.; Plavac, F.; White, C. T. In *Total Synthesis of Natural Products*; ApSimon, J., Ed.; Wiley: New York, 1983; Vol. 5, p 1.
(13) Cane, D. E. In *Biosynthesis of Isoprenoid Compounds*; Porter, J. W., Spurgeon, S. L., Eds.; Wiley: New York, 1981; Vol. 1, p 336.

An Improved IUPAC-Based Method for Identifying Alkanes

SCOTT DAVIDSON*

Computer Data Systems, Inc., One Curie Court, Rockville, Maryland 20850

The IUPAC rules for naming alkanes with more than one main-chain candidate do not always reduce the largest side chains and in some cases fail to break ties, leaving a compound without a unique name. Recognizing that alkane and side-chain structures differ only by level (side chains have main chains that have side chains, etc.) leads to an improved algorithm based upon a simple, recursive definition of a main chain: *the chain with the least complex side chains*. The described computer algorithm breaks down the carbon skeleton into simple and complex side chains, which are coded onto a tree-structured, complexity-ordered specification page suitable for comparison with alternate configurations to minimize side-chain complexity.

INTRODUCTION

Alkane nomeclature in organic chemistry provides the foundation for building sets of rules to describe compounds of increasing complexity. Traditionally, study of alkane nomenclature in college chemistry courses has been limited to finding the longest chain and selecting the correct chain end to label as C-1. This generally suffices to move on to substituted structures. However, there is more to the topic. Consider the alkane shown in Figure 1. This compound has a symmetric, easily visualized structure reflected in the name 3-isopropylpentane. However, according to IUPAC Rule A-2.6(a) (Appendix I), the correct name is 3-ethyl-2-methylpentane. This rule simplifies by breaking the isopropyl group into a methyl group and a segment of an alternate main chain. Similarly, in the Rule 2.6(a) example, a potential

5-carbon side chain is simplified. IUPAC Rules 2.6(a–d) are a sequence of four tiebreaker tests to be employed when alkanes having two or more longest chains are encountered. (The reader is urged to study the examples carefully.)

Main-chain ties are actually a rather common feature of alkanes, for example, every 2-methyl, 3-ethyl, etc. In fact, all alkanes except methane have at least one main chain tie—the same main chain numbered from the other end. This will be shown later to have relevance to the naming of alkanes beyond numbering.

Now consider the substituted decane shown in Figure 2, which is certainly no more complicated than the examples in Appendix I. The C-skeleton shows two candidate main-chain segments (1–4 and 1'–4') branching from C-5, each with methyl and ethyl side chains. The only difference is that the positions of methyl and ethyl are reversed. Now, applying IUPAC Rules 2.6(a–d) yields the following: (a) both have the same number of side chains; (b) both have the same

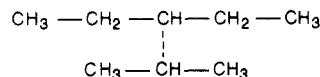* Address correspondence to this author at 240 Manor Circle, No. 2, Takoma Park, MD 20912.
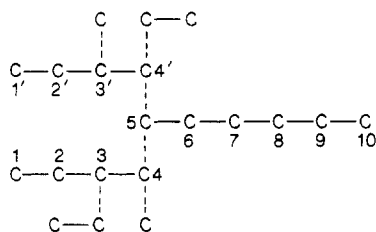
Figure 1. Simplest nonequivalent main-chain tie.

Figure 2. Simplest alkane unnamable by IUPAC rules.

Figure 3. Example of Rule 2.6(a) not reducing largest side chain.

side-chain locants; (c) both have the same number of carbons in the smaller side chains; (d) both have the same degree of branching. The four-step filter has left an unresolved tie and two equally valid names for the alkane:

for 1–4 main chain

3-ethyl-5-(1-ethyl-2-methylbutyl)-4-methyldecane

for 1′–4′ main chain

4-ethyl-5-(2-ethyl-1-methylbutyl)-3-methyldecane

This 20-carbon alkane is the simplest of a class of alkanes that are unnamable by the IUPAC rules. Just two independent entries into a chemical database would have a 50% chance of producing two entries for the same compound. A computer program to identify alkanes by using the IUPAC rules would accept the first-found main chain, which could be either one depending on the structural data input. Nomenclature rules should give unique names. What went wrong here?

Reported deficiencies in the IUPAC rules have generally involved complex ring structures. However, in a 1985 review of organic nomenclature in this journal,[1] Lozac'h gives an example of a chlorinated C_{17} alkane for which naming ambiguities occur, partially due to main-chain ties in the parent hydrocarbon. This example is presented to illustrate the advantages of nodal nomenclature, in which side-chain names are replaced by numeric values of chain lengths and locants. Lozac'h and Goodson[2] have written several papers on applications of this nomenclature system.

IMPROVED ALGORITHM

The main-chain-tie problem originates in Rule 2.6(a). Since alternate main chains have the same length, both have the same number of side-chain carbons (total carbons – main-chain carbons). Therefore, selecting the candidate main chain with the most side chains minimizes the *average* size of all side chains. While this procedure tends to break up large side chains, there is no guarantee that the largest side chain will be simplified as in Figure 1 or that used to illustrate Rule 2.6(a). Consider the next case, which can be handled by Rule 2.6(a) alone, shown in Figure 3.

Rule 2.6(a) requires selection of 1–10 as main chain, forcing the *tert*-butyl group onto a side chain of eight carbons. The alternate selection (placing *tert*-butyl on the main chain 1′–10) would have reduced the largest side chain to six carbons. Now if both chains are extended to the left and right by adding carbons, *tert*-butyl can be replaced by ever larger and complex side chains, while the two methyl groups continue to control selection of the main chain.

The basic flaw in Rule 2.6(a) is that it discards all detailed information about individual side chains. Following a locant rule (having similar problems) in Rule 2.6(b), belated attention is given to side-chain differences in (c) and (d). However, by this time the damage has already been done. More impor-
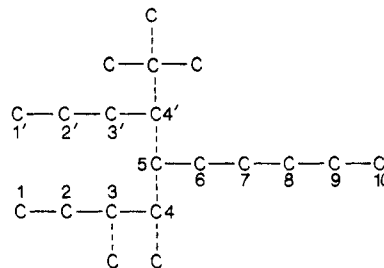
tantly, the detail given to side chains is insufficient, because as structures become larger, side chains begin to resemble general alkanes and require the same detailed scrutiny to identify. An improved method of identifying and naming alkanes based upon this consideration can be formulated in five steps as follows: (1) Recognize the natural tree structure of alkanes. Alkanes and side chains differ only by level; i.e., side chains have main chains that have side chains, etc. (2) Develop a set of rules for measuring side-chain complexity (an algorithm to compare side chains beyond size differences—see Appendix II). (3) Select a set of simple side chains that can be identified without further fragmentation (e.g., the eight C1–C4 alkyl groups). (4) Utilize the recursive relationship

side chain (level) =

main chain (level) + simpler side chains (level + 1)

to simplify structures. (5) Minimize *individual* side-chain complexity by applying the complexity rules consistently at all levels, working upward from the lowest levels.

The above procedure leads to a simple, recursive main-chain definition: *the chain with the least complex side chains*. More specifically, this means the chain whose most complex side chain is least complex. This further implies that the side chains of contending main chains are sorted into descending complexity order for comparison in a tiebreaker series.

Now, just as alkanes have main chains that have side chains (going down-level), by going up-level any alkane can be treated as a side chain of a larger alkane by removing one H from C-1 of the main chain. This yields an alternate main-chain definition: *the chain that minimizes the complexity of the side chain resulting from attaching its C-1 end to a larger alkane.* Both definitions are highly intuitive, as they use the natural tree structure of alkanes—in effect, answering the question, "if the C-skeleton were a tree, which end should be planted in the ground?"

Translation of an alkane structure into a name is actually a recursive or nested procedure for identifying side chains, with second level (and lower) simple alkyl groups placed inside parentheses, as for Figure 2. The process continues until an alkyl main chain is *not* attached to a longer chain. This chain, is, of course, the alkane main chain, and the completion of the process is signaled by replacing the final "yl" with "ane".

COMPLEXITY

The IUPAC 1957 Rules contained a set of rules for determining the relative complexity of alkyl side chains (Appendix II). Former IUPAC Rules A-2.3a(i–iv) were part of an option to order the side chains in the alkane name by increasing complexity rather than alphabetically. They were not employed to identify alkanes. The option was abandoned in the 1971 edition in favor of alphabetical ordering. As a point of interest, Rule 2.3(iii), "The less complex is the chain whose longest substituent has the lower locant", breaks the tie in the "unnamable" alkane of Figure 2 when used in conjunction with the proposed definition "the chain with the least complex side chains". However, these rules have the same drawback as Rules 2.6(a–d) in being inconsistent over levels,
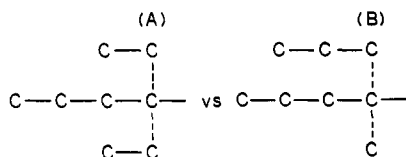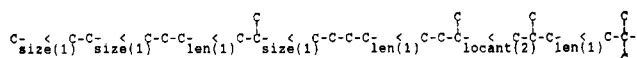
Figure 4. Simplest side chains with the same complexity under the IUPAC 1957 Rules.

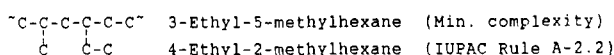yielding unresolved ties. The simplest case is shown in Figure 4.

The intent of the rules seems clear, but some revisions are necessary. First, Rule 2.3(iv) simply implies a tiebreaker series and can be omitted. Rule 2.3(iii) could be modified to read, "The less complex is that whose most complex substituent is least complex". This is, of course, just a restatement of the proposed main-chain definition, except it refers to the entire side chain (or alkane). However, if this rule is changed to, "The less complex of two *identical* substituents is that with the lower locant", then the three tests (size, length, locant) can be applied consistently at each level. In practice this means repeating the size/length tests (going down-level) until branches are small enough to be declared identical by simple tests.

The locant of a side chain is meaningful only as part of a larger structure. Thus, if two isolated side chains being compared for complexity have the same size and length, the 3-step tiebreaker series goes directly to the second level, where the tests are repeated after the branches are sorted into descending complexity order. In Figure 4, (A) is less complex than (B) because ethyl is less complex than propyl at the second level—no locant test is required. [In other words, moving one ethyl group in (A) to the 2-position would create a side chain more complex than (A) but still less complex than (B).] Complexity rankings for the eight designated "simple" (C1–C4) side chains are shown below, along with the test and level required to break the tie:

$$C_{\overline{size(1)}}^{-} < {}_{size(1)}^{C-C} < {}_{len(1)}^{C-C-C} < {}_{size(1)}^{C-C} < {}_{len(1)}^{C-C-C} < {}_{locant(2)}^{C-C-C} < {}_{len(1)}^{C-C-C} < {}^{C-C}_{C}$$

For numeric comparisons, methyl, ethyl, ..., *tert*-butyl can be coded 100, 200, ..., 800. Then their locants can be added as low-order attributes, for example, 3-methyl = 103, 4-ethyl = 204, etc.

By the second (up-level) main-chain definition, an alkane can be evaluated as the least complex of the set of side chains formed by attaching each chain end of the carbon skeleton, one at a time, to a larger ("master") chain at a common locant. Since all side-chain configurations of the alkane have the same size and are considered to have the same locant, the three tests (size, length, locant) reduce to a single test for the simpler alkanes—the familiar longest chain (length) test. Since the set includes both ends of the main chain, selecting one also determines the chain numbering. This process can number oppositely from the IUPAC rules, which are based on locants alone as shown:

```
‾C-C-C-C-C-C‾   3-Ethyl-5-methylhexane   (Min. complexity)
   C̣   C̣-C        4-Ethyl-2-methylhexane   (IUPAC Rule A-2.2)
```

The alkane/chain-end set of side chains can be simulated by attaching *copies* of the C-skeleton to the master chain as isolated side chains, one for each chain end. In the computer simulation shown in Figure 5, the side chains have been sorted by increasing complexity according to the former IUPAC naming option. Thus, the first side chain cited corresponds to the minimum complexity name.

The complexity rules outlined above will now be applied to the examples in Appendix I. While Rule 2.6(a) screens out too many structures by side-chain count alone, the Rule 2.6-(b–d) examples are worth discussion in the context of com-
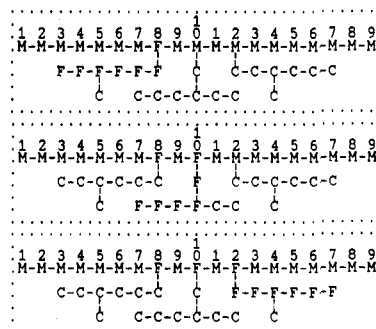
12-(3-METHYLHEXYL)-8-(4-METHYLHEXYL)-10-(2-ETHYLPENTYL)NONADECANE

Figure 5. 3-Methylhexane in three side-chain configurations (M = master chain; F = first-level main chain).

plexity because in two of three cases the main chain with the least complex side chains (as defined above) is not selected. In Rule 2.6(b) the locant test selects the isobutyl side chain over the less complex *sec*-butyl. In Rule 2.6(c) the side chains are ranked in *increasing* complexity order, on the basis of size alone. In this case, the chain with the least complex side chains is selected, but *not* because "2 is greater than 1", as stated in the footnote of Rule 2.6(c), but because 8 is less than 9—reading right to left in decreasing complexity order. (The extra carbon in "2" is already incorporated in "9".) The intent of Rule 2.6(c) is easily defeated by making side chains A and B of Figure 4 segments of candidate main chains. The rule selects main chain A because 2 > 1, although the largest "small" side chain is propyl on main chain B.


```
(A)
   \
    C—C—C—C—C—C      2 2 8  (for A main)
   /
(B)                   1 3 8  (for B main)
```
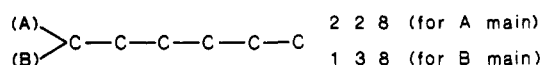
This point is also well illustrated by the final example under Rule 2.6(d). The only choice here is whether to put isopropyl on the main chain or on the large side chain. At first glance, the selection is puzzling because placing isopropyl on the large side chain creates the *most* branched side chain. Then it is realized that the rankings for both choices under Rule 2.6(c) would be 3,8—requiring a tiebreaker. The selection results from comparing propyl vs isopropyl first. However, when the C_8 side chains are evaluated first, isopropyl is seen to be more complex than propyl—at the second level. The minimum complexity name then becomes 5-isopropyl-6-(1-propyl-pentyl)dodecane. By contrast, a different, more complicated example illustrating Rule 2.6(d) in the IUPAC 1957 Rules (Appendix II) was named correctly in terms of minimum complexity and also in the presumed spirit of the rule.

COMPUTER SEARCH/ENCODING

The computer program (FORTRAN, 500 lines) developed to identify alkanes by the new definition utilizes a depth-first search to perform the same sequence of tests at all levels. Side chains are found and characterized in three nested loops over (1) all levels of (2) all ties of (3) all candidate main chains. The search begins at the first chain end in a list of chain ends (equivalently, the first copy of the alkane attached to the master chain). The carbons are counted, and an initial main chain is found. Next, the search moves across the initial main chain, identifying simple side chains. When a complex side chain (>4 carbons) is encountered, the search moves to the next lower level; the locant carbon of the parent main chain now becomes the current chain end, replacing the locant of the hypothetical master chain, and the test sequence is repeated. A local maximum depth of the search is reached when a main chain having only simple side chains is traversed. At this point it becomes feasible to find and break main-chain ties. The side chains of candidate main chains must first be sorted in descending complexity order for comparison. The
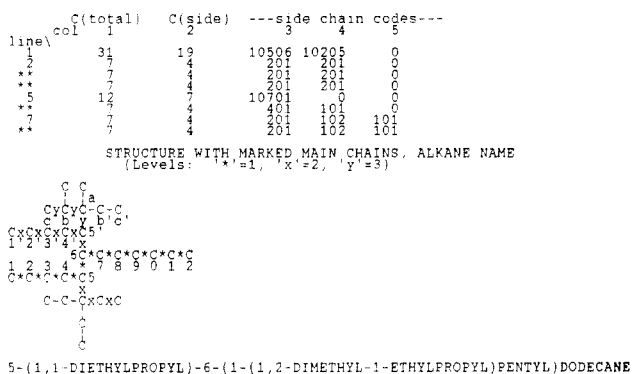
```
              C(total)   C(side)   ---side chain codes---
          col    1          2        3     4     5
     line\
       1         31         19      10506 10205     0
       2          7          4        201   201     0
      **          7          4        201   201     0
      **          7          4        201   201     0
       5         12          7      10701     0      0
      **          7          4        401   101     0
       7          7          4        201   102   101
      **          7          4        201   102   101

          STRUCTURE WITH MARKED MAIN CHAINS, ALKANE NAME
            (Levels: '*'=1, 'x'=2, 'y'=3)

                    C C
                    C a
                CyCyC=C-C
                C'y C y b'c'
            CxCxCxCxC5'
            1'2'3'4'x
                  5C*C*C*C*C*C
            1 2 3 4 *5 7 8 9 0 1 2
            C*C*C*C*C5
                        x
                  C-C-CxCxC
                          ^
                          :
                          C

     5-(1,1-DIETHYLPROPYL)-6-(1-(1,2-DIMETHYL-1-ETHYLPROPYL)PENTYL)DODECANE
```

Figure 6. Main-chain ties at two levels.

search then backtracks to the parent chain and continues in the same manner. (Pseudocode for the above process is given in Appendix III.)

Sorting the side chains to break stored at a given level requires that all lower level ties be broken first so as to specify each side chain in its minimum-complexity configuration. The depth-first search thus guarantees a minimum complexity path from top to bottom levels. The task is complete when all alkane chain ends have been evaluated as C-1 of a candidate main chain (when the master chain has been traversed).

The results of the search are stored as an alkane specification page, where the top line codes the entire alkane and each lower line (if any) represents a different complex side chain. Column 1, column 2, and columns 3, 4 ... quantify the previously stated simplifying formula for side chains:

side chain (level) = main chain (level) + simpler side chains (level + 1)
col 1 (size)   col 2 (size-length)   cols 3,4...(codes)

The length is stored as its complement to permit a continuous low compare when ties are broken. Columns 3, 4 ... hold either a simple side chain code or a complex side chain code that points to the line where the complex side chain is further simplified.

The codes panel of the page is analogous to the lines below, which depict a main computer directory on the top line and subdirectories on lower lines. The side chains of the alkane main chain replace the main directory, complex side chains replace subdirectories, and simple side chains replace filenames.

<MAIN>	file1.top	file2.top	<SUB1>	file3.top
<SUB1>	file1.mid	<SUB2>	file2.mid	file3.mid
<SUB2>	file1.btm	file2.btm		

Listing the files by depth-first search gives the sequence file1.top, file2.top, file1.mid, file1.btm, file2.btm, file2.mid, file3.mid, file3.top. The page is compared, sorted, and finally translated by this method of "reading", whereby the entire page has been read when the top line has been read. (The translation requires fewer than 150 lines of FORTRAN.)

A current best page is initialized as a copy of the first completed page (consisting of the first-found main chain and its side chains) and then updated if improved. Two pages being compared for complexity can also specify two different alkanes. Thus, an alkane database could be ordered by complexity by first translating each alkane name into its carbon skeleton structural formula.

COMPUTER RUN

The edited output of a computer run shown in Figure 6 illustrates the specification page and also why main-chain ties must be broken at the lowest levels first.

Two 12-carbon candidate main chains each have 7-carbon (C_7) side chains of equal length. The asymmetric (upper) side chain has two nonequivalent 3-carbon main-chain candidates. If chain ab'c' is selected, then the most complex side chain

at the third level is isopropyl, vs ethyl for the lower C_7 side chain. This would break the alkane main-chain tie in favor of the 1'–5' branch. However, the upper small side chains can be simplified (ethyl, methyl, methyl) by selecting abc. This is less complex than (ethyl, ethyl) of the lower side chain, making the 1–5 branch a better choice for main chain.

The specification page (Figure 6) contains the following information: Line 1 holds the code for the complete alkane—indistinguishable from a side chain except for being the top line. Code 10506 is for the C_{12} complex side chain comprised of five carbons from the alternate (1'–5') main chain and the upper C_7 side chain. The code indicates that this side chain is further simplified on line 5 and its locant is 6 on the parent chain. Line 2 codes the lower C_7 side chain. Lines 3 and 4 (asterisks) show its equivalent (tie) three-carbon main chains deleted during the search. Line 5 codes the upper C_7 side chain simplified in line 7. Line 6 was its first-found main chain ab'c' subsequently replaced by line 7. Line 8 shows the equivalent chain resulting from the 2-methyl group coded in line 7, column 4.

CONCLUSION

Complexity ordering as utilized to identify a main chain as the chain with the least complex side chains is both more powerful and intuitive than the set of rules in current use and also offers a structurally based alternative to alphabetical ordering for alkane databases.
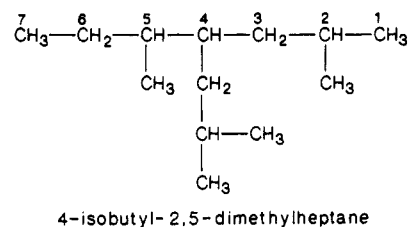
APPENDIX I: IUPAC RULES A-2.6(A–D)[3]

2.6 If chains of equal length are competing for selection as main chain in a saturated branched acyclic hydrocarbon, then the choice goes in series to the following:
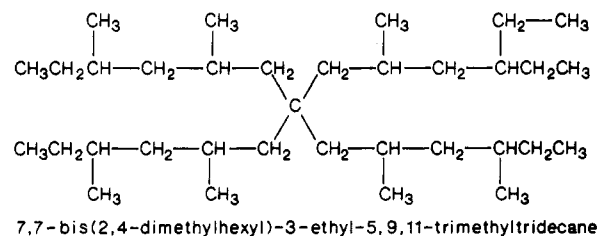
(a) the chain having the greatest number of side chains

2,3,5-trimethyl-4-propylheptane
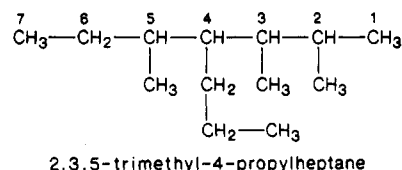
(b) the chain whose side chains have the lowest numbered locants

4-isobutyl-2,5-dimethylheptane

(c) the chain having the greatest number of carbon atoms in the smaller side chains

7,7-bis(2,4-dimethylhexyl)-3-ethyl-5,9,11-trimethyltridecane

Here the choice lies between two possible main chains of equal length, each containing six side chains in the same positions. Listing in increasing order, the number of carbon atoms in the several side chains of the first choice as shown and of the alternate second choice results as

first choice   1, 1, 1, 2, 8, 8

second choice   1, 1, 1, 1, 8, 9

The expression, "the greatest number of carbon atoms in the smaller side chains", is taken to mean the largest side chain at the first point of difference when the size of the side chains is examined step by step. Thus, the selection in this case is made at the fourth step where 2 is greater than 1.

(d) the chain having the least branched side chains

6-(1-isopropylpentyl)-5-propyldodecane
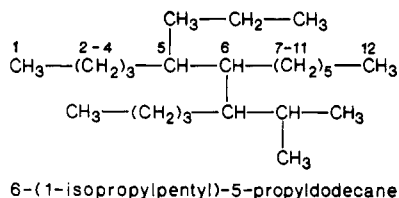
## APPENDIX II: IUPAC 1957 RULES A-2.3A(I–IV) AND 2.6(D) EXAMPLE[4]

2.3. If two or more side chains of different nature are present, they may be cited (a) in order of increasing complexity or (b) in alphabetical order.

(a) The side chains are arranged in order of increasing complexity by applying the following in series until a decision is reached:
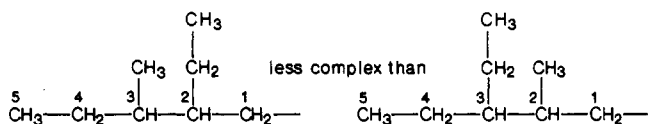
(i) the less complex is that containing the smaller total number of carbon atoms
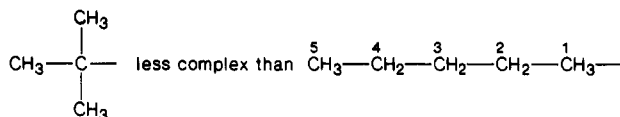
(ii) the less complex is that containing the longer straight chain
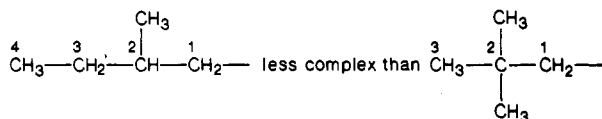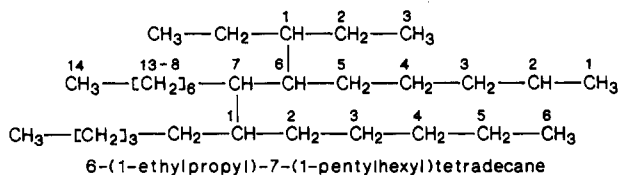

(iii) the less complex is that whose longest substituent has the lower locant

(iv) the less complex is that whose next longest substituent has the lower locant

2.6(d) the chain having the least branched side chains

6-(1-ethylpropyl)-7-(1-pentylhexyl)tetradecane
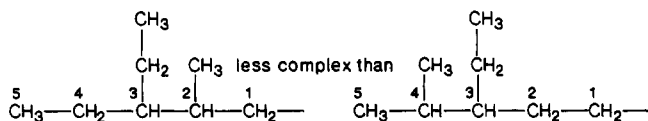
APPENDIX III: PSEUDOCODE FOR ALKANE IDENTIFICATION PROGRAM

Variables shown in uppercase correspond to those actually utilized in the program.

START (at either end of the master chain)
LEVEL=1; LFREE=1; new-side-chain=true (top line of SPEC page)
while LEVEL>0
    TIES=true
    while TIES (>1 main-chain candidates)
        if new-side-chain then (not returning to parent chain)
            LCUR=LFREE (next free line becomes current)
            ICOL=2 (side-chain codes begin col 3 of LCUR)
        more-side-chains=true
        while more-side-chains (step along main chain)
            if new-side-chain found then
                find size, main-chain length and locant
                mark the main chain (e g 'C' → 'X'; prevents revisit, display use)
                ICOL=ICOL+1 (move to next col of LCUR of SPEC page)
                if simple then (C1–C4)
                    enter 3-digit code (KLL, K=1–8, LL=locant)
                else (complex)
                    LFREE=LFREE+1 (allocate next free line)
                    enter 5-digit code (1NNLL, NN=LFREE, LL=locant)
                    SPEC(1,LFREE)=size; SPEC(2,LFREE)=size-length (no. of side C's)
                    store search-parameters(LEVEL) (to backtrack)
                    LEVEL=LEVEL+1
                    MAIN1(LEVEL)=true (1st-found longest chain)
                    LCUR=LFREE; ICOL=2
                    new-side-chain=true
                    start at "C0" of side chain = current locant of main
            else more-side-chains=false
        end(More-side-chains)
        sort down side-chain codes on LCUR (+ lines linked by complex codes)
        if not MAIN1 then (evaluating tie)
            compare LCUR with previous best: if less complex then change NN of parent-line code to LCUR
        resume search for next tie main chain from where current was found
        if found then
            marked tie main-chain replaces current
            MAIN1(LEVEL)=false
            LFREE=LFREE+1 (allocate new line for tie-spec)
        else
            unmark current main chain ('X' → 'C'; free for revisit as side chain)
            TIES=false
    end(TIES)
    LEVEL=LEVEL-1
    if LEVEL>0 then recall search-parameters(LEVEL); new-side-chain=false
    (return to "C0" of parent chain; parent line/col of SPEC page)
end(LEVEL>0)
DONE (can translate SPEC page into alkane name)

REFERENCES

(1) Lozac'h, N. Principles for the Continuing Development of Organic Nomenclature. *J. Chem. Inf. Comput. Sci.* **1985,** *25,* 180–5.
(2) Goodson, A.; Lozac'h, N. Advantages of Nodal Numbering for Uniquely Identifying Atoms in Chemical Nomenclature. *Croat. Chem. Acta* **1986,** *59*(3), 547–63.
(3) IUPAC. *Nomenclature of Organic Compounds*; Pergamon Press: New York, 1979; pp 10–11.
(4) IUPAC. *Nomenclature of Organic Compounds*; Butterworth: London, 1958; pp 8–9.