the Third International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems; July 1990; ACM Press: Charleston, SC, 1990; pp 1073–1079.
(2) Bleyberg, M. Z.; Zhou, T.; Isenhour, T. L.; Marshall, J. C. An Expert System Model for Analytical Chemistry. Int. J. Appl. Artif. Intell., accepted for publication.
(3) The Development of a Robotic Standard Addition Method. Eckert-Tilotta, S. E.; Isenhour, T. L.; Marshall, J. C. Anal. Chim. Acta, accepted for publication.
(4) The Application of Standard Robotic Methods for Water Analysis. Lee, J. R.; Isenhour, T. L.; Marshall, J. C. J. Chem. Inf. Comput. Sci. 1991, 31, 000–000.
(5) Wirfs-Brock, R.; Wilkerson, B.; Wiener, L. Designing Object-Oriented Software; Prentice Hall: New York, 1990.
(6) Tello, E. Object-Oriented Programming for Artificial Intelligence; A Guide to Tools and System Design; Addison-Wesley Publishing Co.: Reading, MA, 1989.
(7) Smith, M.; Tockey, S. An Integrated Approach to Software Requirements Definition Using Objects; Boeing Commercial Airplane Support Division: Seattle, WA, 1988; p 132.
(8) Shank, R.; Riesback, C. R. Inside Computer Understanding; Lawrence Erlbaum Associates: Hillsdale, NJ, 1981.

Wiener and Randič Topological Indices for Graphs

BERTHOLD J. MAIER

IDC Internationale Dokumentationsgesellschaft für Chemie, Otto-Volger-Strasse 19, W-6231 Sulzbach, Germany

Received June 27, 1991

We present counterexamples to two conjectures on the monotony of the Wiener and Randič indices for graphs raised in a recent paper,[1] and we prove that the conjecture on the Randič index holds for graphs with vertex degrees bounded by 8 and a strengthening holds for the chemically relevant graphs with vertex degrees bounded by 4.

INTRODUCTION

For an unoriented connected graph **G** with vertex set V, the Wiener index[2] χ_W is defined by

$$\chi_W(G) = \sum d(v,v')/2$$

where the sum is taken over all vertex pairs $v,v' \in V$, and $d(v,v')$ denotes the distance, i.e., the length of the shortest path, between v and v' in **G**.

The Randič index[3] χ_R is defined for an unoriented graph **G** by

$$\chi_R(G) = \sum 1/\sqrt{\deg(v)\,\deg(v')}$$

where the sum is taken over all edges vv' of **G**, and $\deg(v)$ denotes the vertex degree of v in **G**.

These are two of the many indices[4,5] which have been proposed in order to correlate macroscopic properties with graph theoretical invariants of the structural formula of the organic molecule. In a recent paper[1] the following conjectures were raised:

$$\chi_W(G') < \chi_W(G) \tag{1}$$

$$\chi_R(G') < \chi_R(G) \tag{2}$$

where **G** is a connected graph with $n \geq 2$ vertices and **G′** is the subgraph obtained from **G** by deleting the last vertex in a canonical labeling of **G**. In ref 1 a labeling of **G** is called canonical if (a) it is a cooperative labeling,[1,6] i.e., starting with a vertex v_0 we next label the neighbors of v_0 in a given order, then the neighbors of the neighbors in the chosen order and so on and (b) the lower triangular submatrix of the adjacency matrix is lexicographically maximal when its rows are concatenated to a string. The conjectures 1 and 2 have been checked for all graphs with up to 9 vertices.[1]

In this paper we give counterexamples to both conjectures, but we show that conjecture 2 holds for graphs with vertex degrees bounded by 8 and a strengthening of conjecture 2 is true for graphs with vertex degrees bounded by 4. Our examples show that these bounds are optimal. The counterexample for conjecture 1 is a cycle with vertex degrees equal to 2.

THE WIENER INDEX

Let P_n denote a straight chain with n vertices and C_n a simple cycle with n vertices. Clearly, $C_n' = P_{n-1}$ in the notation of the conjecture. Now, using $1 + 2 + ... + i = \binom{i+1}{2}$ we obtain

$$\chi_W(P_n) = \sum_{i=2}^{n}\binom{i}{2} = \binom{n+1}{3} \tag{3}$$

$$\chi_W(C_{2n}) = n\left[\binom{n}{2} + \binom{n+1}{2}\right] = n^3 \tag{4}$$

$$\chi_W(C_{2n+1}) = (2n+1)\binom{n+1}{2} \tag{5}$$

The following gives the values of χ_W for some of these graphs which disprove conjecture 1:

G	P_9	P_{10}	P_{11}	C_{10}	C_{11}	C_{12}
$\chi_W(G)$	120	165	220	125	165	216

THE RANDIČ INDEX

We start with an example that was also considered independently by S. Tratch.[7] Let G_n denote the graph consisting of two bridge-heads A,B joined by n chains with two vertices each. Thus, G_n has $2n + 2$ vertices, the vertices A,B have degree n, and the other vertices have degree 2. In a canonical labeling one bridge-head, say A, has label 1, and the other bridge-head, B, has the last label. This holds, because after A we first label all the neighbors of A, the first elements in the chains, then the second elements in the chains, and finally we label B as the last element. Therefore, $G_n' = G_n - \{B\}$.

Now B is connected in G_n to n elements of degree 2, which each are connected to the second element of degree 2 in the respective chains. In G_n' the elements to which B is connected in G_n all have degree 1. Therefore, we have the contribution $n/\sqrt{2}$ from the chain bonds to $\chi_R(G_n')$ in G_n'. In G_n we have the contribution $n/2$ to $\chi_R(G_n)$ from the chain bonds, and $n/\sqrt{2n}$ from the bonds incident with B. The bonds incident with A give the same contributions to $\chi_R(G_n')$ and $\chi_R(G_n)$. Hence conjecture 2 holds for G_n if and only if

$$n/\sqrt{2} < n/2 + n/\sqrt{2n} \tag{6}$$

But this is false for $n \geq 6 + 4\sqrt{2}$.

Thus, G_{12} is a counterexample to conjecture 2. The graph G_{12} has 2 vertices of degree 12 and, therefore, is of theoretical interest only.

We proceed with two other examples which will show the optimality of the positive results. Let H_n be the rooted tree whose root has n descendents each of which has $n-1$ leaves. Let H_n'' denote the graph obtained from H_n by omitting the root. Then $\chi_R(H_n'') < \chi_R(H_n)$ if and only if

$$(n-1)/\sqrt{n-1} < (n-1)/\sqrt{n} + 1/n \qquad (7)$$

which is false for $n \geq 5$.

We now start with two copies of the rooted tree H_n and let I_n denote a graph consisting of the two copies and $n(n-1)$ edges corresponding to a bijection between the sets of leaves of the two rooted trees. As for G_n it is easily seen that a canonical labeling of I_n starts with one of the roots and ends with the other. Let I_n' denote the graph obtained from I_n by removing one of the roots. It follows that $\chi_R(I_n') < \chi_R(I_n)$ if and only if

$$(n-1)/\sqrt{2(n-1)} < (n-1)/\sqrt{2n} + 1/n \qquad (8)$$

which is false for $n \geq 9$. Hence I_9 is a graph with vertex degrees bounded by 9, which violates conjecture 2. In the following we will show that conjecture 2 holds for all graphs with vertex degrees bounded by 8. We begin with a result with less restrictions on the removed vertex, which yields a strengthening of conjecture 2 for the chemically relevant graphs with vertex degrees bounded by 4.

Theorem 1. Let G be a graph and G″ a subgraph obtained from G by deleting a nonisolated vertex of degree at most 4 and its incident edges. Then

$$\chi_R(G'') < \chi_R(G)$$

Corollary 1. Let G be a graph with vertex degrees bounded by 4 and G″ a subgraph obtained from G by deleting a nonisolated vertex and its incident edges. Then

$$\chi_R(G'') < \chi_R(G)$$

The graph H_5 shows that the bound on the degree cannot be improved without further restrictions.

Proof: Let A denote the vertex deleted from G to obtain G″. By assumption, A has degree $a = 1, 2, 3,$ or 4. We consider the neighbors of A in G and the contributions of the edges incident with these neighbors to $\chi_R(G'')$ and $\chi_R(G)$, respectively. Edges incident with exactly one neighbor of A are associated with this neighbor. We proceed to associate the edges incident with two neighbors of A.

(*1*) Suppose BC is an edge between two neighbors B,C of A, i.e., A is contained in the triangle ABC in G. If B has a degree of 2 in G, then we associate BC with B. Let c denote the degree of C in G″. For the edges incident with B we have for BC the contribution $1/\sqrt{c}$ to $\chi_R(G'')$ and $1/\sqrt{2(c+1)}$ to $\chi_R(G)$. For the edge AB we have the contribution $1/\sqrt{2a}$ to $\chi_R(G)$. For $2 \leq c$ and $a \leq 4$ we have for the contribution of BC to $\chi_R(G'')-\chi_R(G)$ in comparison with the contribution of AB to $\chi_R(G)$

$$1/\sqrt{c} - 1/\sqrt{2(c+1)} < 1/\sqrt{8} \leq 1/\sqrt{2a} \qquad (9)$$

For $c = 1$ we get $1/2 < 2/\sqrt{8} \leq 2/\sqrt{2a}$ with the additional contribution of AC to $\chi_R(G)$. In the following we only need to consider edges between such neighbors of A whose vertex degrees in G″ are at least 2.

(*2*) Suppose BC is such an edge between two neighbors B,C of A, both of whose degrees b and c, respectively, are at least 2 in G″. The edge BC contributes $1/\sqrt{bc}$ to $\chi_R(G'')$ and $1/\sqrt{(b+1)(c+1)}$ to $\chi_R(G)$. We observe that $1/\sqrt{1+x} > 1 - x/2$, $x \neq 0$, the right-hand side being the tangent line to the left-hand side at $x = 0$. We obtain

$$1/\sqrt{b+1} = 1/\sqrt{1+1/b}\cdot 1/\sqrt{b} > [1 - 1/(2b)]/\sqrt{b} \qquad (10)$$

and

$$1/\sqrt{(b+1)(c+1)} > [1 - 1/(2b)][1 - 1/(2c)]/\sqrt{bc} \text{ if } > [1 - 1/(2b) - 1/(2c)]/\sqrt{bc} \qquad (11)$$

For the contribution of BC to $\chi_R(G'')-\chi_R(G)$ we obtain the bound

$$1/\sqrt{bc} - 1/\sqrt{(b+1)(c+1)} < [1/(2b) + 1/(2c)]/\sqrt{bc} \qquad (12)$$

We now associate $1/(2b\sqrt{bc})$ of this bound to vertex B and $1/(2c\sqrt{bc})$ to vertex C.

(*3*) Let B be a neighbor of A with degree b in G″. Suppose there are b' edges incident with B and no other neighbor of A. These edges contribute $s'/\sqrt{b}$ to $\chi_R(G'')$ and $s'/\sqrt{b+1}$ to $\chi_R(G)$, where $s' = \sum_{i=1}^{b'} 1/\sqrt{d_i}$ and d_i are the degrees of those neighbors of B in G″ and in G which are not neighbors of A. Hence the contribution of these edges to $\chi_R(G'')-\chi_R(G)$ is

$$s'(1/\sqrt{b} - 1/\sqrt{b+1}) \qquad (13)$$

Suppose next that there are b'' edges incident with B and other neighbors C_i of A with degrees $c_i \geq 2$ in G″. Clearly, $b' + b'' \leq b$. In the previous paragraph we have associated $s''/(2b\sqrt{b})$ of the bounds for the contributions of these edges to $\chi_R(G'')-\chi_R(G)$ to the vertex B, where $s'' = \sum_{i=1}^{b''} 1/\sqrt{c_i}$, and the $c_i \geq 2$ are the degrees of the neighbors C_i of A incident with an edge BC_i. We have not yet taken into account the contribution $1/\sqrt{a(b+1)}$ of the edge AB to $\chi_R(G)$. The theorem will be proved if we can show that

$$s'(1/\sqrt{b} - 1/\sqrt{b+1}) + s''/(2b\sqrt{b}) < 1/(2\sqrt{b+1}) \text{ if } \leq 1/\sqrt{a(b+1)} \qquad (14)$$

for $a \leq 4$, which indicates that the contribution of the edge AB to $\chi_R(G)$ is greater than the contributions of the edges in G″ to $\chi_R(G'')-\chi_R(G)$. For this it suffices to show

$$2s'(\sqrt{1+1/b} - 1) + s''\sqrt{b+1}\big/(b\sqrt{b}) < 1 \qquad (15)$$

(*4*) Observe that $\sqrt{1+x} < 1 + x/2$, $x \neq 0$, the right-hand side being the tangent line to the left-hand side at $x = 0$. Since $s' \leq b'$ we obtain for the first summand in eq 15:

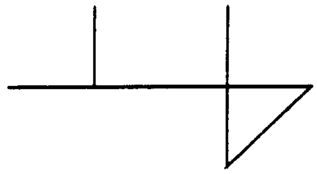
$$2s'(\sqrt{1+1/b} - 1) < s'/b \leq b'/b \qquad (16)$$

(*5*) For the second term in eq 15 we observe that $s'' \leq b''/\sqrt{2}$, since $c_i \geq 2$ for all i. Note that $\sqrt{b+1}/\sqrt{2b} < 1$, since $b + 1 < 2b$ for $b \geq 2$. Hence we obtain for the second summand in eq 15

$$s''\sqrt{b+1}\big/(b\sqrt{b}) < b''\sqrt{b+1}\big/(b\sqrt{2b}) \leq b''/b \qquad (17)$$

Now eq 15 follows from $b' + b'' \leq b$, and the theorem is proved.

We return to conjecture 2, but we also prove a result for removing the last vertex in a broader class than canonical labelings. We call a cooperative labeling a cooperative degree-decreasing labeling, abbreviated to cdd, if the neighbors of each vertex are labeled with nonincreasing degrees. Here is a graph whose canonical labelings are not cdd:



Theorem 2. Let G be a graph with a cdd or a canonical labeling whose last vertex has a degree from 1 to 8, and let G′ be the subgraph obtained from G by deleting this last vertex. Then

$$\chi_R(G') < \chi_R(G)$$

Corollary 2. Let G be a graph with vertex degrees bounded by 8, and let G′ be the subgraph obtained from G by deleting the last vertex of a cdd or a canonical labeling of G. Then

$$\chi_R(G') < \chi_R(G)$$

The graph I_9 shows that the bound on the degree is optimal for the given conditions.

Proof: (*1*) Let A be the last vertex in G, and let A have degree $a = 1, ..., 8$. As in the proof of theorem 1, we consider the contribution of the edges incident with neighbors of A to $\chi_R(G')-\chi_R(G)$. As before, an edge incident with only one neighbor of A is associated with this neighbor. An edge BC between two neighbors B,C of A with degrees *b*,*c* gives the contribution $1/\sqrt{bc} - 1/\sqrt{(b+1)(c+1)}$ to $\chi_R(G')-\chi_R(G)$, and this is bounded from above by $1/(2b\sqrt{bc}) + 1/(2c\sqrt{bc})$. If both $b,c \geq 4$ then we associate $1/(2b\sqrt{bc})$ of this contribution to vertex B and $1/(2c\sqrt{bc})$ to vertex C. These contributions are bounded by $1/(4b\sqrt{b})$ for B and by $1/(4c\sqrt{c})$ for C. If $b \leq 3$ then we associate

$$1/\sqrt{bc} - 1/\sqrt{(b+1)(c+1)} - 1/(4c\sqrt{c}) \qquad (18)$$

to B and $1/(4c\sqrt{c})$ to C. If both $b,c \leq 3$ then we make a choice which vertex gets the contribution according to eq 18.

(*2*) Suppose that a neighbor B of A has a neighbor D of degree 1. If the labeling is cdd then also A has degree 1, since A appears after D in the labeling. If the labeling is canonical then A has degree 1 by the maximality of the labeling. In both cases the theorem, therefore, follows from theorem 1. In the following we can assume that all neighbors of neighbors of A have a degree of at least 2 in G.

(*3*) Suppose that a neighbor B of A is connected only to other neighbors of A apart from A itself. Then it follows from cdd or from the canonical labeling that the degree of A is at most the degree of B. In particular, if the degree of B in G′ is at most 3, then we can omit these cases by reference to theorem 1.

(*4*) Now consider a neighbor B of A with degree *b* in G′ which only gets contributions of the form $1/4b\sqrt{b}$ from edges with other neighbors of A. This holds whenever $b \geq 4$ or $b \leq 3$ and no edge with another neighbor of A is attributed to B according to eq 18. Assume there are b' edges between B and nonneighbors of A with degrees $d_1, ..., d_{b'}$. The contribution of these edges to $\chi_R(G')-\chi_R(G)$ is $s'(1/\sqrt{b} - 1/\sqrt{b+1})$ with $s' = \sum_{i=1}^{b'} 1/\sqrt{d_i} \leq b'/\sqrt{2}$, since $d_i \geq 2$ by point 2. Combining this with $1/\sqrt{b} - 1/\sqrt{b+1} < 1/(2b\sqrt{b+1})$ we obtain the upper bound

$$b'/[b\sqrt{8(b+1)}] \qquad (19)$$

for the contribution of the edges to nonneighbors of B to $\chi_R(G')-\chi_R(G)$.

Assume that there are b'' edges between B and other neighbors of A. The contribution of these edges attributed to B is bounded from above by

$$b''/(4b\sqrt{b}) \qquad (20)$$

Taking the two bounds together we have

$$b'/[b\sqrt{8(b+1)}] + b''/(4b\sqrt{b}) \qquad (21)$$

and the theorem holds if this is less than the contribution of the edge AB to $\chi_R(G)$, namely

$$1/\sqrt{a(b+1)} \geq 1/\sqrt{8(b+1)} \qquad (21a)$$

for $a \leq 8$. For this it suffices that

$$b'/b + b''\sqrt{b+1}\big/(b\sqrt{2b}) \leq 1 \qquad (22)$$

which is obvious from $b' + b'' = b$ and $b + 1 \leq 2b$, $b \geq 1$.

(*5*) Finally we consider the remaining cases where B gets a contribution of the form (eq 18) from an edge with another neighbor C of A. Suppose B and C have degrees *b*,*c* in G′, respectively. By construction $b = 1, 2,$ or 3. Let b'' be the number of edges between B and other neighbors of A. Since $b \leq 3$ the cases $b'' = b$ have been dealt with in point 3, and we are left with the cases $b = 2$, $b'' = 1$ and $b = 3$, $b'' = 1, 2$.

For the contribution to $\chi_R(G')-\chi_R(G)$ of an edge BD with a nonneighbor D of A with degree *d* we have

$$1/\sqrt{bd} - 1/\sqrt{(b+1)d} \leq 1/\sqrt{2b} - 1/\sqrt{2(b+1)} \qquad (23)$$

for $d \geq 2$ by point 2.

We first consider the case $b = 2$, $b'' = 1$. We have one edge with a neighbor C of A of degree *c* in G′ and one edge with a nonneighbor D. The edge AB contributes $1/\sqrt{a(b+1)} \geq 1/\sqrt{24}$ to $\chi_R(G)$. For $\chi_R(G')-\chi_R(G)$ we have the upper bound

$$(1/\sqrt{2c} - 1/\sqrt{3(c+1)} - 1/(4c\sqrt{c})) + (1/\sqrt{4} - 1/\sqrt{6}) - 1/\sqrt{24} \qquad (24)$$

In order to simplify the dependence on *c* we use

$$1/\sqrt{c+1} > [1 - 1/(2c)]/\sqrt{c} \qquad (25)$$

which is eq 10 with *b* replaced by *c*. We then obtain the monotone upper bound

$$(1/\sqrt{2} - 1/\sqrt{3})/\sqrt{c} + (1/\sqrt{12} - 1/4)/(c\sqrt{c}) + (1/\sqrt{4} - 1/\sqrt{6}) - 1/\sqrt{24} \qquad (26)$$

whose value $1 - 3\sqrt{6}/8 - \sqrt{2}/16 \approx -0.0069$ for $c = 2$ is negative. Recall that the case $c = 1$ was covered in point 3.

Next consider the case $b = 3$, $b'' = 1$. We have one edge to a neighbor C and two edges to nonneighbors. The edge AB contributes $1/\sqrt{a(b+1)} \geq 1/\sqrt{32}$ to $\chi_R(G)$. For $\chi_R(G')-\chi_R(G)$ we have the upper bound

$$[1/\sqrt{3c} - 1/\sqrt{4(c+1)} - 1/(4c\sqrt{c})] + 2(1/\sqrt{6} - 1/\sqrt{8}) - 1/\sqrt{32} \qquad (27)$$

which is less than zero, as we will see in the last case.

Finally we handle the case $b = 3$, $b'' = 2$ with two edges with neighbors of A and one edge with a nonneighbor. We may have one or two of the bonds to neighbors attributed to B according to eq 18. In the first case let C be the neighbor whose edge is handled according to eq 18. From the other edge with a neighbor we have the contribution $1/(4b\sqrt{b}) = 1/(12\sqrt{3})$. For $\chi_R(G')-\chi_R(G)$ we have the upper bound

$$[1/\sqrt{3c} - 1/\sqrt{4(c+1)} - 1/(4c\sqrt{c})] + 1/(12\sqrt{3}) + (1/\sqrt{6} - 1/\sqrt{8}) - 1/\sqrt{32} \qquad (28)$$

which is less than zero, as we will see in the next case.

In the last subcase with two edges with neighbors attributed to B according to eq 18, we show that half of the contribution of the edge AB and the edge to the nonneighbor cancel with each of the contributions according to eq 18. Thus,

$$(1/\sqrt{3c} - 1/\sqrt{4(c+1)} - 1/(4c\sqrt{c})) + (1/\sqrt{6} - 1/\sqrt{8})/2 - 1/(2\sqrt{32}) \quad (29)$$

is an upper bound for one half of the contribution of the edges at B to $\chi_R(G')-\chi_R(G)$. We use eq 25 once more and obtain the monotone upper bound

$$(1/\sqrt{3} - 1/2)/\sqrt{c} + (1/\sqrt{6} - 1/\sqrt{8})/2 - 1/(2\sqrt{32}) \quad (30)$$

whose value $\sqrt{6}/4 - 7\sqrt{2}/16 \approx -0.0063$ for $c = 2$ is negative. Since the constants $2(1/\sqrt{6} - 1/\sqrt{8}) - 1/\sqrt{32} \approx -0.067$ in eq 27 and $1/(12\sqrt{3}) + (1/\sqrt{6} - 1/\sqrt{8}) - 1/\sqrt{32} \approx -0.074$ in eq 28 are less than the constant $(1/\sqrt{6} - 1/\sqrt{8})/2 - 1/(2\sqrt{32}) \approx -0.063$ in eq 29, we have shown that $\chi_R(G')-\chi_R(G)$ is negative in all cases, and the proof of the theorem is complete.

REFERENCES AND NOTES

(1) Kvasnička, V.; Pospíchal, J. Canonical Indexing and Constructive Enumeration of Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 99–105.
(2) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
(3) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
(4) Rouvray, D. H. The Modelling of Chemical Phenomena Using Topological Indices. *J. Comput. Chem.* **1987**, *8*, 470–480.
(5) Rouvray, D. H. The Limits of Applicability of Topological Indices. *THEOCHEM* **1989**, *54*, 187–201.
(6) Morgan, H. L. The Generation of a Unique Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **1965**, *5*, 105–113.
(7) Pospíchal, J. Private communication, 1991.

Domain-Oriented Knowledge-Acquisition Tool for Protein Purification Planning

HENRIK ERIKSSON

Department of Computer and Information Science, Linköping University, S-581 83 Linköping, Sweden

Received August 22, 1991

Knowledge acquisition for expert systems can be facilitated by appropriate supporting tools. By providing knowledge-acquisition tools that closely resemble the experts' own conceptual model of the domain, the knowledge-acquisition bottleneck can be widened significantly. P10 is a domain-oriented knowledge-acquisition tool specialized to the problem area of protein purification planning. P10 enables specialists in protein purification to develop expert systems for this task without a mediating knowledge engineer. Knowledge structures entered in a graphical interface are transformed into appropriate knowledge base constructs by P10.

1. INTRODUCTION

Knowledge acquisition (KA) has long been recognized as a major bottleneck in the development of expert systems. Computer-based tools that supports different aspects of the KA process have been suggested and, to a varying degree, implemented and used for practical purposes. However, little has been published on KA tool support for chemical informatics. Commercial expert systems shells offer a level of abstraction (e.g., rules and objects) suitable for knowledge engineers, but current shells are largely unsuitable for domain experts.

P10 is a knowledge-acquisition tool that provides specialists in protein purification planning with means to enter, edit, and review their planning knowledge according to a conceptual domain model supported by the tool. The user is required to fill in forms and to enter *partial plans*[1] (i.e., planning schemata or recipes) in a graphical user interface. The knowledge acquired by P10 is automatically transformed through a knowledge base generator into knowledge bases that can be used immediately for testing. Thus, experts can use P10 to develop the bulk of an expert system without an intermediate knowledge engineer.

Earlier KA tools have been based on a particular knowledge elicitation technique (e.g., ETS[2]) or a model of the problem-solving method used (e.g., MORE,[3] MOLE,[4] and SALT[5]). Knowledge acquisition tools such as P10 draw their power from an abstract model of the domain in question rather than other (more general) models such as models of classification, planning, and design. The major advantages of this approach are that specialized tool support is provided, which typically leads to better understanding of and acceptance for the KA tool among experts, and that KA environments can be tailored to the requirements from a small group of experts. The rest of this article develops the idea of domain-oriented KA tools for protein purification and reports the domain model supported by P10 as well as the P10 implementation.

2. BACKGROUND

Research in biochemistry requires isolation of proteins that are present in only trace amounts from biological material. Such an isolation is performed in a number of stages of which purification by high-resolution chromatography techniques is the last one. The outcome of the purification depends on the order in which different chromatography techniques are applied and on the running conditions for each of them. A complete purification plan is more or less impossible to construct in advance, since the effect of each operation cannot be fully predicted (at least when important properties of the protein is unknown). Thus, planning protein purification requires experience and skill.

2.1. Planners for Protein Purification. In an earlier project we have studied computer-supported planning for protein purification. As part of this project an expert system (P8) that supports a chemist in planning protein purification was implemented. P8 supports liquid chromatography techniques applied to membrane-bound proteins. The system gives advice on which techniques to use, the running conditions for each step, and measurements to be taken between each purification step. The P8 planner is based on a set of *partial plans*,[1] i.e., skeletal solutions that could be refined and combined into a

0095-2338/92/1632-0090$03.00/0