

高分子の C-13 NMR 化学シフトエキス パートシステム

巽 久行・西岡 篤夫

An Expert System of C-13 NMR Chemical Shift of Polymers

Hisayuki TATSUMI and Atsuo NISHIOKA

Synopsis

An expert system (ES) of Carbon-13 NMR Chemical shift of polymers has been developed by the use of a commercially available tool, SOGEN, for personal computer (e.g. PC9801). This system can display a list of some plausible names of synthetic polymers among selected 14 kinds of polyolefins including some copolymers by the input of any set of the chemical shifts of polymers, observed or trially selected values in the range from 7 to 50 ppm. This range was adopted to satisfy the upper limit of the number of the chemical shift values with 1 ppm interval. The certainty factor (CF) was assumed to distribute over ± 1 ppm around the referenced observed values of the chemical shifts of each polymer with triangle membership function based on fuzzy method. By setting up 44 rules, which have integer values of chemical shift with 1 ppm interval in LHS, and possible polymer names with their calculated CF values in RHS, and further introducing some correction to the CF values for large discrepancies from expected output, good results has been obtained for the system in serving to predict possible polymer names with descending order of their CF values.

概 要

高分子の Carbon-13 NMR 化学シフトを入力して高分子名を推論するエキスパートシステムを市販のパソコン用ツール「創玄」を用いて作成した。事象変数の選択枝数の制限から対象とする高分子は共重合体を含めてポリオレフィンのみとし、14 種を選んだ。化学シフトは 1 ppm 間隔で 7-50 ppm の範囲とし、基準とする各試料のシフト値の確信度係数 (CF) の計算には三角形分布を仮定し、設定値の左右 1 ppm の幅をとった。1 ppm 間隔の各シフトを条件部とし、その区間にある複数の高分子毎に CF の計算値を与える方法でルールを 44 個作成した。推論結果が著しく期待に反する場合はルールの結論部に必要な修正を加えた。これにより 14 種の高分子につきほぼ満足すべき結果が得られた。

1. 緒 言

高分子の核磁気共鳴 (NMR) スペクトルは高分子合成並びに物性の分野で重要な研究手段として多用されているが、特に Carbon-13 (C13) NMR スペクトルは高分子化合物の主要な構成原子である炭素に着目して分子構造情報を得るもので最も重要な分析法の一つとなっている。同時に未知化合物のスペクトルデータから化合物名の帰属を効率よく行うことも必要であり、このためデータブックやデータベースが利用されている。他方近年知識ベースと推論エンジンを組合わせたエキスパートシステム (ES) の研究開発が盛んに行われ¹⁾、パソコンで動くツールもかなり市販されるに至った。

著者は以前から高分子の C13 NMR 化学シフトデータベースの構築^{2,3)}を行っており、この共同利用も実現したが、化合物名や構造に関する情報から数値データを検索する機能とは逆に本研究では化学シフトを入力してその帰属を行う作業をパソコン上で ES に

より実現することを試みたのでこれについて報告する。

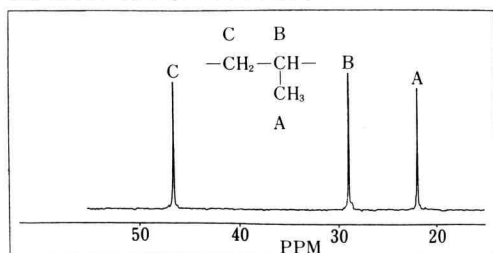
2. 高分子の C13 NMR 化学シフト

はじめに高分子の C13 NMR 化学シフトと分子構造解析の関係について略述する。

通常高分子化合物を構成する原子の種類は炭素、水素、酸素、窒素等が主なものであるが、C13 NMR はこのうち炭素の原子核、特に質量数 13 の同位体が磁気能率を持ち核磁気共鳴 (NMR) を示す事に着目し、かつこの共鳴核のスペクトル位置 (化学シフト) が隣接する原子との結合状態や分子間相互作用により微妙に変化することから、化学シフト値と分子構造との関係が生ずるのである。図 1 は代表的な合成高分子の一つであるポリプロピレン (polypropylene) の C13 NMR スペクトルであるが、その分子構造 (CH₂CHCH₃) に対応して左から CH₂, CH, CH₃ の各炭素核のシグナルが観測される。この様な C13 NMR スペクトルは分子の種類だけでなく、構成単位である単量体 (モノマー) の連鎖の規則性その他の微細構造の影響も受けるので、シフトの値や数も変化する。低分子化合物に比べて高分子化合物はスペクトルに与える要因が多いので分子構造の精密な分析が一般に容易ではないが、C13 NMR 法は概して尖鋭なスペクトルを示すので情報量が多く、分析に多用される所以である。

高分子化合物の研究を行う場合、構造のある程度分ったものについて微細構造を求める事と未知化合物についてそのスペクトルから物質名を同定することは方向が逆である。本研究は後者に関しエキスパートシステムの構築によりこれを支援しようとするものである。

128. ポリプロピレン (アイソタクチック)



溶媒: ODCB, 温度: 135°C, 周波数: 50.1 MHz, 積算: 3,400 回
信号 (PPM) A: 21.84, B: 28.83, C: 46.44

図 1. ポリプロピレンの C-13 核磁気共鳴スペクトル⁴⁾

3. エキスパートシステム構築ツール「創玄」⁵⁾

本研究では将来大規模なシステムを構築するための前段階として先ずパソコン上で動く小規模の ES 作成を目的として市販のツール「創玄」を用いたのでこの機能を簡単に説明する。

最近のツールは知識の表現にルールとフレームを併用するものが多いが、創玄はルールベースを基本とし、後向き推論を採用しており、構造が簡単の上マンマシンインタフェースが比較的良好でできているので小規模 ES の作成に向いている。ルールベースといっても OPS5 流のプロダクションシステムではなく、AND OR TREE 型で分析型 ES に向く。単なる事実の表現にはフレームではなく、事象変数とその選択肢 (値) を組合わせる方法を採用し、ルールの条件部と結論部にこの事象変数を確信度係数 (CF) と組合わせて書くようになっている。また、数値の大小関係を利用したルール表現に事象変数の代わりに数値変数を用いることができる。ES 構築の手順を次に示す。

この ES は条件部に化学シフトの値を入力して結論部に高分子名を書くルール表現が基本になるので、事象変数には化学シフト (shift) と高分子名 (polymer) を採用する。

- 1) 最初にゴールとなる事象変数を決める。ここでは polymer とした。
- 2) ゴール変数の選択肢を入力する。(63 個以内)
- 3) 他の事象変数とその選択肢を入力する。ここでは shift。

- 4) ルールを作成する。条件部、結論部とも複数の選択肢を AND, OR で組合わせることができる。

この場合いずれも確信度係数 (CF) を入力する。

以上で ES 構築の入力作業は終わり、次に実行をメニューから選ぶ。

- 5) ゴールでない条件部の事象変数の選択肢を画面から選ぶ。

- 6) 各選択肢ごとに CF を入力する。(この値が結論に影響する)

- 7) 入力が終わると推論して結論が CF 付きで表示される。

- 8) ファンクションキーの HOW 機能で推論の過程を知ることができる。

このほかにも説明機能がある。

創玄の主機能は以上の如くで、詳細は後述する。

4. 化学シフトの原データ

本研究では高分子の C13 NMR スペクトルの原データとして測定条件が同じで、スペクトルの帰属が行われており、かつ代表的な試料が含まれている「高分子分析ハンドブック」⁴⁾のスペクトル図を参照した。これには合成高分子として代表的なポリオレフィン、ビニールポリマー、ジエンポリマー、重付加及び重縮合ポリマー等につき共重合体も含めて 68 図が収録されている。化学シフトの測定精度は 0.01 ppm で、参照したシフト値はポリオレフィン 14 種につき、帰属がなされている約 100 個の値を用いた。引用した高分子試料名と化学シフト数の対応を表 1 に示す。

5. エキスパートシステムの設計

使用したパソコンは NEC の PC-9801 VX で、この上で「創玄」を動かした。これは前述の如く事実とルールとの二つの知識表現により知識を入力し、後向き推論で結論を得る仕組みになっている。入力上の制限はルールは数百個程度で、事象変数は 255 個以内であるが、各変数毎の選択肢数は 63 個以内なので設計に若干工夫が必要である。

高分子の C13 NMR 化学シフトの範囲は通常 6-180

ppm であるが、シフト値を事象変数にすると選択肢数を 60 個として 3 ppm 間隔になり粗すぎる。そこで 1 ppm 間隔にすると 50 ppm の範囲しかとれないが、ポリオレフィンに限定し、しかもビニールを含む共重合体を除けばその化学シフトは 7-50 ppm の範囲にあるので、今回は前掲の 68 図から C と H のみから成る 14 種のポリオレフィン試料を選んで化学シフト ES を作成することを試みた。

5.1 事象変数

事象変数は「A は B である」という事実を A を変数名とし、B を選択肢(値)とする知識表現である。ここでは前述の如く、高分子名と化学シフトの二つを事象変数に選んだ。結論は高分子名で出力されるので、ゴール事象変数 polymer の選択肢として表 1 に示す 14 個を用いた。化学シフト事象変数の選択肢は 7-50 ppm の範囲を 1 ppm 間隔に 44 個とった。

各高分子の原スペクトル図から帰属のなされているシフト値の数が表 1 で個数として示されている。文献等ではもっと詳細な帰属の報告もあるが、ここでは前記ハンドブックのデータのみを用いた。また事象変数 polymer の選択肢として通常用いられている慣用名は長すぎるので、表 1 に示す略語を使用した。但しこの略語はすべて一般に用いられているわけではない。

表 1. ポリオレフィンの名称(慣用名)と選択肢名(略語)及びシフトの個数

高分子名	選択肢名	シフトの個数
高密度ポリエチレン	HDPE	4
低密度ポリエチレン	LDPE	13
エチレン-プロピレン共重合体	CEP2 (c2 rich)	5
エチレン-ブテン-1 共重合体	CBUTE	7
エチレン-ヘキセン-1 共重合体	CEHXE	9
エチレン-4-メチルペンテン-1 共重合体	CEPNE	8
エチレン-オクテン-1 共重合体	CEOCE	9
ポリプロピレン (アイソタクチック)	IPP	3
ポリプロピレン (シンジオタクチック)	SPP	5
エチレン-プロピレン ランダム共重合体	CEP3 (C3 rich)	8
エチレン-プロピレン ブロック共重合体	CEPB (block)	10
プロピレン-ブテン ランダム共重合体	CBUTP	8
ポリブテン-1	PBUT	4
ポリ-4-メチルペンテン-1	PPNE	5

5.2 ルールの表現と信頼度係数 (CF: Certainty Factor)

高分子名とシフトを対応させるには二つの方法が先ず考えられる。一つは前者を条件部に与えて後者を結論部を書くもので、他はその逆である。高分子名をゴールにしたのでここでは後者を採用した。条件部には AND/OR を用いて複数の条件を書くことができ、結論部には OR で複数の結論を書ける。

シフトを条件部を書く場合でも、1 個の polymer を念頭において完全な帰属に必要な複数の値を入力する方法と 1 ppm 幅に入る複数の polymer を念頭において 1 個のシフトを条件部にかき、可能性のある複数の polymer を結論部を書く方法とがある。前者は入力シフトが 100% 条件部に一致しない限り hit が 0 になるので好ましくない。後者は推論に弾力性があるが、仕組みが複雑になりルール数が多くなるとシステム構築者の負担が大きくなるという問題もある。ここでは後者の方法によった。

確信度係数 (CF) はルールベースの硬直性を緩和する機能があるが、その値の設定は ES の実用性の観点から経験的になり易く、システムの保守を面倒にする欠点もある。本研究では化学シフト ES の実用性から CF を用いるが、主観的な設定を抑制する為 CF の入力につきファジー推論のメンバーシップ関数にならって次に述べる一定の基準で計算した値を用いることにした。

5.3 CF の評価方法

シフト間隔は 1 ppm であり、一方原データは 0.01 ppm の精度で与えられている。ユーザが選択するのは画面の 7, 8, 9, …等のシフト値であるから、原データとのずれを CF に反映させる必要がある。CF には一般的に考えると着目した炭素種の化学シフトの他にスペクトルの強度や幅その他の要素が関係すると考えられるが、ここでは単純にシフトのみに CF を結びつけたルールを作ることにした。

原データの CF の分布は三角形と仮定し、中心値の左右 1 ppm に広がるものとする。例えば 25.68 ppm というシフトは 25.68 ppm にピークを持ち、24.68-26.68 にまたがることになり、全体は 24-25, 25-26, 26-27 ppm の 3 区間となる。従って各区間での CF は次のようにして計算される。三角形の面積を 1 とすると

$$24-25 \text{ ppm} : CF = (1-0.68)(1-0.68)/2 \\ = 0.0512$$

$$25-26 \text{ ppm} : CF = (0.5-0.0512) + (0.5-0.2312) \\ = 0.7176$$

$$26-27 \text{ ppm} : CF = 0.68 \times 0.68/2 \\ = 0.2312$$

原データが丁度 25.00 ppm の場合は 24-25, 25-26 ppm の両区間でいずれも CF は 0.5 となるし、また 25.50 ppm の時は上の 3 区間で夫々 0.125, 0.75, 0.125 となり、中央の区間で最大になる。このことは入力画面で選択肢である 1 ppm 間隔のシフトを選ぶ時、暗にその中央を期待したことと同様であり、厳密にはその妥当性に問題もあるが測定条件や試料によるシフトの変動が通常 0.1 ppm 以上あることから第 1 近似として差支えないと考えられる。

この方法で原図にあるシフト値からその左右を含む 2-3 区間での CF を計算し各ルールの結論部に与える候補 polymer の CF 値としたが、創玄の実行結果をみると次に述べる複数のルールによる CF 値の相乗効果によりシフトの個数 (ルール数) が多くなるとゴール時の CF が急速に 1 に近づき、多数候補 polymer の優劣差が僅少で判定困難となることが分った。そこでこの判定を容易にするため CF の三角形分布の面積を以下の計算では 0.5 とした。次の例はポリプロピレン (PP) の代表的シフトの一つである CH 基のもつ 28 ppm 帯に着目したルールであるが、この前後の 27 及び 29 ppm 帯に値を持つ polymer も CF に寄与するので結論部はかなり複雑になっている。参考のため各 polymer のシフトを付記した。

例：もし The chemical shift is 28 ppm ならば

The polymer is

IPP	(28.83 ppm)	(確信度 = +0.32)
CEP3	(28.74)	+0.38
CBUTP	(28.25, 28.74)	+0.36
CEPB	(27.48, 28.83)	+0.25
PBUT	(27.78)	+0.16
SPP	(28.48)	+0.07
CEHXE	(27.33, 29.57)	+0.08
CEP2	(27.47)	+0.07
CBUTE	(27.35)	+0.03
LDPE	(27.33)	+0.03
CEOCE	(27.28, 27.33)	+0.03

5.4 CF 値の組合わせ計算

創玄では複数のルールが適用される時、途中のルールによる CF の相互作用を考慮したければ各 CF を組合わせて計算することができる。この結果、ルールが多い程最後の CF 値は高くなる。例えば 2 個のルールがそれぞれ CF1, CF2 を持つとき、推論が終わると結果の CF は次式で与えられる。

$$CF = CF1 + (1 - CF1)CF2$$

3 個以上では $CF + (1 - CF)CF3$ となり、以下同様で残りの CF 値即ち $(1 - CF)$ に対して後のルールの CF が乗ぜられるので、初期に 1 に近づくと後のルールは CF に殆ど効果がない。但しノイズと考えられる候補を抑制するためには後述の様に $CF = -1$ を与えることも必要になる。

6. ES 構築過程の諸問題と修正作業

前項で述べた設計方法により実際に ES の構築作業を進めてみると、推論結果が期待するものとかかなり異なる場合（ノイズ出力）が生じた。これにはルールの結論部の表現が不十分なことによる場合と、ルールそのものに別種の工夫を要する場合とがあるが、試行錯誤的な修正作業が必要になる。実用性を目的とする ES にとってこの作業はある程度不可避なことであると思われる。

6.1 $CF = -1$ の併用によるノイズの抑制

例えば CEP2 と CBUTE は前者の 20 ppm の信号を除けばかなり似ており、しかも前者のシフトを全部入力しても後者の 30 ppm 帯の 2 個のシフトによる CF の相乗効果のため殆ど同順位となり経験的な推論とは矛盾する。この点を改善するため 19 ppm, 20 ppm のルールの結論部に CBUTE の $CF = -1$ を加えた結果、CEP2 のシフトのみの入力に対して CBUTE は出力されなくなった。多少試行錯誤的であるが、推論結果中に出現を予想し難い polymer があり、しかもその CF が高い場合にこれを抑制するのにこの方法は有効である。今回は CF 値の分布をシフト位置の左右 1 ppm 幅で三角形にとったので 1 回でもヒットすれば低い CF でも最後まで残ることがある。

6.2 中間仮説の採用 (PE 及び PP の場合)

HDPE は 30 ppm に著しく強い信号がある他は微弱的な 3 個があるだけであるが、LDPE には分岐による弱い信号が多く、原図で帰属されているものだけで 13 個もある。両者を区別し易くするため事象変数として「Other shift」を追加し、選択肢として存在の可否を問うことにした。結論の候補に HDPE がヒットしたときこの事象変数と HDPE を組合わせたルールの結論部の CF は HDPE のみ 1 で他はすべて -1 にすると HDPE 以外のノイズは殆ど消去された。PP の場合も立体規則性によりアイソタクチック、シンジオタクチック、アタクチック等の異性体があるので中間仮説の採用をした。この方法は中間段階で高い CF を得たノイズを抑制する効果がある。

7. 実行結果

以上の修正作業を行って構築した化学シフト ES につき、原図のシフトを整数値にして入力した結果の推論出力順位は以下の如くである。

7.1 HDPE

4 個のシフト値の小数部を切捨て、Other shift not exist と入力すると

HDPE	CF = +0.94
CEOCE	CF = +0.52

となり、また小数部を四捨五入し、Other shift not exist と入力すると

HDPE	CF = +0.92
CEOCE	CF = +0.51
LDPE	CF = +0.10

となり、妥当な結果が得られる。CEOCE の出現は HDPE の 4 個のシフトをすべて含むためである。

7.2 LDPE

13 個のシフトを全て入力すると下記の推論結果が出力される（下欄 CF）。

LDPE	CBUTP	CEPB	PBUT	SPP
1.00	0.58	0.58	0.53	0.38

LDPE の CF が 1 になるのは 7 ppm, 8 ppm にキーとなるシフトがあり、このとき LDPE のみ $CF = 1$ とし、

他はすべて-1としたためである。但しこの信号は弱くて観測されない場合があるのでこれを除いて入力しても下記の如く LDPE が首位であることに変わりはない。

LDPE	CEPB	CEP3
0.97	0.77	0.58
CBUTP	PBUT	SPP
0.58	0.53	0.38

7.3 PP

21, 28, 46 ppm の 3 個を入力し, Other shift not exist により下記出力となる。

PP	IPP	CEPB	CBUTE	PPNE
0.75	0.75	0.69	0.35	0.34
CEHXE	LDPE	CEOCE		
0.08	0.03	0.03		

Other shift exist と入力すると SPP (シンジオタクチック) が首位となる。

SPP	PP	CEPB	PPNE	CEP3
0.92	0.61	0.58	0.34	0.23
CBUTP	LDPE			
0.05	0.03			

PP が同時に出るのは Other shift を含むルールが PP を条件部に持つからである。

7.4 CEP

エチレン-プロピレン共重合体は通常はランダム共重合体でエチレン (C₂H₄) が多いもの (CEP2), プロピレン (C₃H₆) が多いもの (CEP3) とがあるが, この他に両者がブロックになっているもの (CEPB) があり, 各モノマーの組成比と結合の仕方によってシフトが若干変化する。各試料のデータを全て整数値で入力した結果を以下に示す。

CEP2		CEP3	
(20, 27, 30, 33, 37)		(20, 21, 24, 28, 37, 46)	
CEP2	0.88	CEP3	0.91
CEPB	0.83	IPP	0.77
LDPE	0.63	PP	0.77
CEP3	0.57	PPNE	0.53
SPP	0.43	LDPE	0.41
PBUT	0.34		
CBUTP	0.07		

PPNE	0.07
CEPB	
(21, 27, 28, 30, 33, 37, 38, 46)	
CEPB	0.96
CEOCE	0.81
CEHXE	0.81
IPP	0.74
PP	0.74
CEPNE	0.68
LDPE	0.64
CEP3	0.57
PPNE	0.39
CEP2	0.38
CBUTP	0.11

CF の計算に用いた各試料のシフトの分布を比較してみると下記の如くである。

シフト	20	21	22	23	24	25	26	27	28	29
CEP2	20.0							27.6		
CEP3	20.8	21.8			24.6				28.7	
CEPB		21.8						27.5	28.8	
IPP		21.8							28.8	
SPP	20.3	20.8							28.5	
シフト	30	31	32	33	34	35	36	37	38	39
CEP2	30.0				33.3			37.6		
CEP3								37.8		
CEPB	30.0	30.4	30.9	33.2				37.6	38.0	
IPP										
SPP										
シフト	40	44	45	46	47					
CEP2										
CEP3			46.0	46.4						
CEPB			46.4							
IPP			46.4							
SPP			46.6	47.2						

CEP3 と CEPB はシフトがよく似ているが 24 ppm の有無がキーであることが分るから 24 ppm のルールに CEPB の CF=-1 を使うと両者の CF 値に格差を与えることができるが, CEP2 との区別はまだ十分ではない。

7.5 PBUT と CBUTE

両者のシフトを比較すると

PBUT 10.7 27.8 35.2 40.4
CBUTE 11.1 26.8 27.4 30.0 30.5 34.1 39.8

30 ppm の有無がキーであることが分る。30 ppm のルールに PBUT の CF=-1 をつかうと次の様に妥当な結果が得られる。シフトの個数は夫々4, 7である。

PBUT	0.83	CBUTE	0.97
CBUTP	0.67	PBUT	0.58
CBUTE	0.64	CBUTP	0.57

推論を終了しました。次に、結論を表示します。

THE POLYMER IS	(1.01) PBUT	(確信度 = +0.83)
	(1.02) CBUTP	(確信度 = +0.67)
	(1.03) CBUTE	(確信度 = +0.64)
	(1.04) SPP	(確信度 = +0.38)
	(1.05) CEP(block C3 rich)	(確信度 = +0.25)

図2. 結論表示画面の例

「THE POLYMER IS PBUT」(確信度 = +0.83)ということは 4 個のルールを適用することにより示されました。
そのうちで 1 番目に適用されたルールを表示します。

以下のルールを適用することにより(確信度 = +0.34)になりました。

ルール 2 1
もし
1) THE CHEMICAL SHIFT IS 27 PPM (+0.20<=確信度<=+1.00)
ならば
2) THE POLYMER IS CBUTE (確信度 = +0.14)
3) THE POLYMER IS CBUTE (確信度 = +0.37)
4) THE POLYMER IS CEP (確信度 = +0.38)
5) THE POLYMER IS CBUTP (確信度 = +0.07)
6) THE POLYMER IS LDPE (確信度 = +0.35)
7) THE POLYMER IS PBUT (確信度 = +0.34)
8) THE POLYMER IS CEHXE (確信度 = +0.35)
9) THE POLYMER IS CEPNE (確信度 = +0.30)
10) THE POLYMER IS CEOCE (確信度 = +0.35)
11) THE POLYMER IS SPP (確信度 = +0.38)
12) THE POLYMER IS CEP(block C3 rich) (確信度 = +0.25)
13) THE POLYMER IS HDPE (確信度 = -1.00)

図3. HOW 機能による適用ルールの追跡(その1)

「THE POLYMER IS PBUT」(確信度 = +0.83)ということは 4 個のルールを適用することにより示されました。
そのうちで 2 番目に適用されたルールを表示します。

以下のルールを適用することにより(確信度 = +0.58)になりました。

ルール 4
もし
1) THE CHEMICAL SHIFT IS 10 PPM (+0.20<=確信度<=+1.00)
ならば
2) THE POLYMER IS PBUT (確信度 = +0.35)
3) THE POLYMER IS CBUTE (確信度 = +0.20)
4) THE POLYMER IS CBUTP (確信度 = +0.20)
5) THE POLYMER IS CEP (確信度 = -1.00)
6) THE POLYMER IS LDPE (確信度 = +0.25)
7) THE POLYMER IS CEHXE (確信度 = -1.00)
8) THE POLYMER IS CEPNE (確信度 = -1.00)
9) THE POLYMER IS CEOCE (確信度 = -1.00)
10) THE POLYMER IS PP (確信度 = -1.00)

図4. HOW 機能による適用ルールの追跡(その2)

「THE POLYMER IS PBUT」(確信度 = +0.83)ということは 4 個のルールを適用することにより示されました。
そのうちで 3 番目に適用されたルールを表示します。

以下のルールを適用することにより(確信度 = +0.74)になりました。

ルール 3 4
もし
1) THE CHEMICAL SHIFT IS 40 PPM (+0.20<=確信度<=+1.00)
ならば
2) THE POLYMER IS PBUT (確信度 = +0.37)
3) THE POLYMER IS CBUTE (確信度 = +0.14)
4) THE POLYMER IS CBUTP (確信度 = +0.34)
5) THE POLYMER IS LDPE (確信度 = -1.00)
6) THE POLYMER IS PP (確信度 = -1.00)

図5. HOW 機能による適用ルールの追跡(その3)

「THE POLYMER IS PBUT」(確信度 = +0.83)ということは 4 個のルールを適用することにより示されました。
そのうちで 4 番目に適用されたルールを表示します。

以下のルールを適用することにより(確信度 = +0.83)になりました。

ルール 29

もし

ならば 1) THE CHEMICAL SHIFT IS 35 PPM (+0.20<=確信度<=+1.00)

2) THE POLYMER IS PBUT (確信度 = +0.34)
3) THE POLYMER IS CBUTP (確信度 = +0.30)
4) THE POLYMER IS CEPNE (確信度 = +0.20)
5) THE POLYMER IS CEPNE (確信度 = +0.25)
6) THE POLYMER IS CEOCE (確信度 = +0.09)
7) THE POLYMER IS PP (確信度 = -1.00)

図 6. HOW 機能による適用ルールの追跡(その 4)

候補 CBUTP		候補 PPNE		候補 CEPNE		候補 CEHXE		候補 CEOCE	
CBUTP	0.96	PPNE	0.80	CEPNE	0.97	CEHXE	0.99	CEOCE	0.98
CEP3	0.73	CEPNE	0.54	CEP	0.61	CEOCE	0.97	LDPE	0.95
CEPB	0.69	CEHXE	0.38	PPNE	0.58	LDPE	0.93	CEPB	0.77
CBUTE	0.62	PP	0.38	CEPB	0.58	CEPB	0.75	CEP2	0.66
SPP	0.47	CEPB	0.38	CEP3	0.39	CEP2	0.66	CEP3	0.55

SPP 0.38 CEPB 0.54
CEPB 0.25 SPP 0.38
CEP3 0.34
PPNE 0.31

化学シフトの入力による推論結果の順位と CF 値が次々と求められてゆく過程は HOW 機能を用いて追跡することができる。図 2~6 は化学シフトとして 10, 27, 35, 40 の 4 個を入力した場合の結果で PBUT が一位であるが、これについて HOW 機能でルール群がどのように使われ、CF がどのように加算されてゆくかがよく分る。この機能はルールの内容や CF 値の調整を行うのにも役立つ。

7.6 その他

以上の他 CBUTP, PPNE, CEPNE, CEHXE, CEOCE などについて得られた推論結果から上位各 5 個を示すと次の通りで、各出力リスト中の首位はいずれも候補名と一致している。下位の CF 値の比較については今後検討の要があるが、第一近似としてはほぼ満足できるものである。

8. 結 論

合成高分子の Carbon-13 NMR 化学シフトのエキ

スパートシステム(ES)構築の第一段階として今回 14 種のポリオレフィンに関して帰属のなされたシフト値を基準値にとり、これと対応する高分子名との関係を市販のパソコン用ツール「創玄」を用いてルール化することを試みた。この場合確信度係数(CF)計算はファジー流に三角形のメンバーシップ関数を仮定して行った。システムの整合性をチェックするため原データを整数値にして 1 ppm 間隔で刻まれた 7-50 ppm の範囲で入力した結果、複数ルールの適用により計算された CF 値の出力順位は首位についてはいずれも予想と一致し、シミュレーションとして一応満足すべきものと思われるが、以下に述べるいくつかの問題が残されており、今後の研究により本システムの有効性を更に高めるように努力する考えである。

1) C13 NMR スペクトルのデータには化学シフトのほか信号強度、幅、精度その他があるが、今回は試行的な意味で最も簡単な場合として化学シフト値のみを用いてルールベースを作成したので、推論結果は CF に大きく依存し、分布関数の形と広がりシフト幅の選び方が問題となる。幅を狭くするとヒットは少なくなるが、可能性があっても全く出てこなくなり融通性がない。幅を広くするとこの反対になり一長一短である。今回は ± 1 ppm の幅にとったが、今後尚検討の余地がある。

2) 予想外のノイズを消すため必要に応じて

CF=-1を採用したが、ルール数が多くなると全体としての調整作業が繁雑になるという問題もある。シフト間隔を適宜不定長にしてルール数が過大にならぬ様にするのも一法であろう。

3) ルールの作成については、化学シフトを数値変数として狭い範囲毎に高分子名を出力する形式も考えられ、入力値と基準値との差から CF をリアルタイムで計算できれば実験者にとってもっと身近なシステムになろう。

4) 強度の導入は当然行うべきであるが、ルール作成が必ずしも容易ではない。強度の設定レベル数でルールが倍増することも考えなければならない。

5) 「創玄」の上位版「大創玄」では事象変数の選択肢数の上限(63)が除かれたので、180 ppm の全範囲を扱うことができ、また事実の表現に事象変数のほかフレームによる知識表現も可能なので化学シフト以外

の要因の導入も含めて今後検討の予定である。

終わりに本研究は昭和 62 年度から 3 年間文部省科学研究費補助金の交付を受けて行われたものでここに感謝の意を表す。

参 考 文 献

- 1) 上野晴樹, 知識工学入門, オーム社, 昭和 60 年
- 2) 西岡篤夫, 京都大学大型計算機センター広報, 16 巻, No. 5, 253-265 (1983)
- 3) 西岡篤夫, 畑田耕一, 藤原 譲, 末広祥二, 同上, 19 巻, No. 2, 54-56 (1986)
- 4) 日本分析化学会編, 高分子分析ハンドブック, 朝倉書店, 昭和 63 年
- 5) AI ソフト, 「創玄」操作マニュアル, 昭和 63 年