

コリレーションによる

5L-3

ハイパーテキストと関係データベースの統合

王忠清\* 山口和紀\*\* 北川博之\*\* 大保信夫\*\* 藤原譲\*\*  
 (筑波大学工学研究科\* 筑波大学電子情報工学系\*\*)

1. はじめに

近年、大量文献情報のハイパーテキスト化の要求が高まっている。この目的のためにはハイパーテキストにおけるブラウジングの技術とデータベースにおける大量情報の検索技術を統合することが有力である。ここでは、有機合成に関する個々の論文から構成される"Organic Synthesis"ジャーナルを対象としたブラウジングシステムにおける両技術の統合方式について述べる。

2. 本研究の目的

大量の文献集合から構成される本研究の対象をハイパーテキスト化する場合、一つの文献内でネットワークを構築しブラウジングすることに関しては、従来のハイパーテキスト技術をそのまま応用すれば充分である。

しかしながら、文献間に存在する関連に関しては、一つの文献がそれ自身の内容が閉じているため内部でのノード間の関連が強いのに比し、それほど強いものとは言えず、目的に応じて航行の方向が多様に変化する傾向がある点、文献が極めて大量である点から、従来のハイパーテキストにおけるリンクで表現すると(1)オーサリングおよび(2)航行の制御が極めて困難であるという問題点が生ずる。

この問題の解決法として対象とされる文献集合中に現われる各種の事実を整理した表、各用語の間の一般的な知識を表わした表等をデータベース化し、文献内の各ノードと上の表中の各値の対応を用意することにより、文献の与えられたノードからデータベース内の知識を介して、他の文献のノードへ航行する方法を提案する。具体的にはわれわれはハイパーテキストとデータベースを結びつける役割を果たすコリレーションを導入し、このコリレーションを扱うことにより実現する。こうするとアーティクル間の関連を実現できるし、また、問い合わせで航行するために航行の向きが柔軟になる。

本システムは三つの部分から構成される：ハイパーテキスト部分、関係データベース部分と両者の対応関係を表わすコリレーション部分である。

3. システムの構成

(1)ハイパーテキスト部

ハイパーテキスト部はノードとリンクから構成される。ノードはテキストやイメージなどの情報が格納される。リンクはノード間の関係を表すものである。

既に述べたように、本研究の対象の各文献是一群のノードとリンクからハイパーテキストとして格納し、それぞれをアーティクルと呼ぶ。アーティクルはarid(article-id)により識別される。アーティクル内の各ノードにはローカルに決められたlnid(local node-id)が与えられる。arid+lnidによって、各ノードがユニークに識別される。またリンクにも同じようにローカルに決められたllid(local link-id)が与えられ、arid+llidによって、ユニークに識別される。

(2) 関係データベース部

文献集合にある各種事実を整理した表、各用語の間の一般的な知識を表わす表などの情報をここに格納する。

例えば：

Markovnikov水和

オレフィン	反応時間	生成物	収率(%)	TID
1-ヘキセン	10分	2-ヘキサノール	96	#1
1-ドデセン	70分	2-ドデカノール	91	#2
2-メチル-1-ブテン	5分	2-メチル-2-ブタノール	90	#3
2-メチル-2-ブテン	10分	2-メチル-2-ブタノール	95	#4

TIDはタプル-id#を表わすものである。

(3) コリレーション

コリレーションは関係データベース世界とハイパーテキスト世界との対応関係を表すリレーションである。

コリレーションのスキーマは次のように定義される。

correlation(arid, lnid, rid, aid, tid)

arid : アーティクルのid

lnid : ローカルなノードのid

rid : リレーションのid

aid : attributeのid

tid : タプルのid

これによって、関係データベースにあるリレーションのvalue(rid, aidおよびtidで決める)とハイパーテキストにあるnode(aridおよびlnidで決める

間の対応関係を表示できる。たとえば下の例ではMarkovnikov水和リレーションにある2-メチル-1-ブテンと2-メチル-2-ブテンという二つのvalueがそれぞれハイパーテキストの(#a1, #110)と(#a3, #112)で表現する二つのノードと対応していることを表わしている。

コリレーションの例:

arid	lnid	rid	aid	tid
#a1	#110	Markovnikov水和	オレフィン	#3
#a3	#112	Markovnikov水和	オレフィン	#4

(4) 例

関係データベースを利用することにより、たとえば今、表示している生成物(アークティクル#a1のノード#110)を、2倍以下の反応時間内により良い収率で生成するオレフィンを探すというような問い合わせは次のように実現できる。問い合わせはSQL風に記述する。

external \$REL, \$ANAME, \$TNUM (\$REL, \$ANAME と\$TNUMはそれぞれリレーション名、attribute名とタプルのidを表わす外部変数である。)

```
SELECT ($REL,rid), ($ANAME,aid), ($TNUM,tid)
FROM correlation
WHERE arid = #a1
AND lnid = #110
```

この問い合わせは図1のAからBをたどる部分に相当する。複数の結果がある場合はそれぞれに対して次の問い合わせを行う。

```
SELECT correlation.arid, correlation.lnid
FROM correlation, $REL t, $REL s
WHERE t.TID = $TNUM
AND t.収率 < s.収率
AND t.生成物 = s.生成物
AND t.反応時間 <= s.反応時間*2
AND correlation.rid = $REL
AND correlation.aid = $ANAME
AND correlation.tid = s.tid
```

この問い合わせは図1のBからC経てDに至る部分と相当する。

図1のAからDへのたどり方は、特定の問い合わせに対してはハイパーテキスト中のリンクで表現す

ることができるが、問い合わせをさまざまにかえてたどることはハイパーテキストのリンクではできない。

4. 結論

有機合成に関する文献情報のブラウジングシステムを対象として、コリレーションによるハイパーテキストと関係データベースの統合の方法について提案し、その有効性を示した。今後は、他の事例について、更に検討を進めるとともにプロトタイプの実現を進めていく予定である。

参考文献

- [1] 石塚英弘他 "データ管理機能を強化したHyper Card システム" 情報学基礎13-7, (1989)
- [2] Jeff Conklin, "Hypertext: An introduction and survey," Computer, 20(9), p.17-41(1987)
- [3] 中田貴士他 "ハイパテキストとデータベースの統合化" 情報処理学会第40回(平成2年前期)全国大会

Markovnikov水和

オレフィン	反応時間	生成物	収率(%)	TID
1-ヘキセン	10分	2-ヘキサノール	96	#1
1-ドデセン	70分	2-ドデカノール	91	#2
2-メチル-1-ブテン	5分	2-メチル-2-ブタノール	90	#3
2-メチル-2-ブテン	10分	2-メチル-2-ブタノール	95	#4

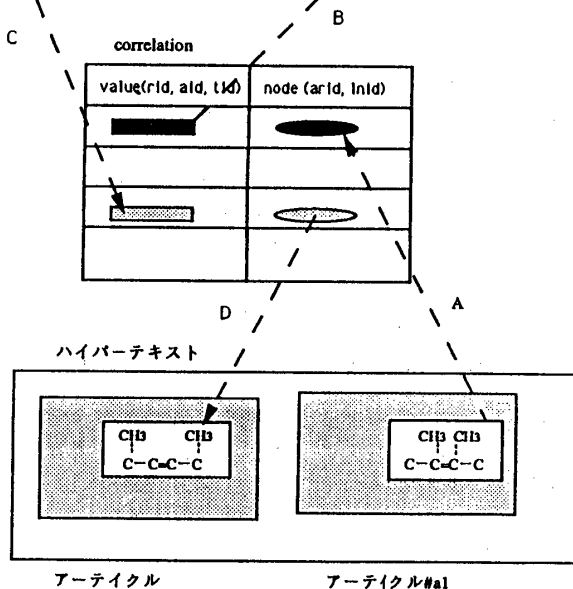


図1 コリレーションの役割