

2Q-8 辞書ハイパーテキスト Hydra(2) — データモデル — \*

根本 治朗, 山下 真司, 松山 洋一, 内藤 広志, 柵木 孝一†  
 キヤノン(株) 情報システム研究所‡

1 はじめに

ハイパーテキストシステムは、柔軟・広範な検索能力を持つDBMSとしての機能を有している。そこで、これまでに蓄積してきたデータをハイパーテキストシステムに格納し、有効利用可能とすることが期待される。ただし、ハイパーテキストシステムのナビゲーションやフィルタリング等の機能を利用するためには、データを構造化する必要がある。そして特に大量のデータを構造化する場合、あらかじめ原文書のデータ構造やそこに含まれている関係を分析する必要がある。

辞書ハイパーテキストHydraの開発の際も、原文書である辞書の構造を調べ、それをモデル化した。このとき用いたダイアグラムは、まず辞書の分析に必要なリンクの分類を行ない、これに基づいてCoadらのダイアグラム<sup>[1]</sup>を拡張したものである。本稿ではモデル化に用いたリンクの分類及びダイアグラムについて述べる。

2 リンクの種類

ハイパーテキストデータは、文書の断片であるノードを、その間に存在している関係を表すリンクによって結びつけたネットワーク構造で文書を表示したものである。しかし、実際の文章では各断片間に様々な関係が存在しているため、単一のリンクだけで文書の構造をすべて表現することは困難である。

これまでもいくつかの文献<sup>[2][3][4]</sup>でリンクの分類が取り上げられているが、それらの多くはリンクの構造や意味だけでなく、その機能やインプリメントの仕方までを含む分類を行なっている。

しかし、例えばリンクの実装法はプログラマの裁量の範囲であり、分析時にそれを決定する必要はない。そこで我々はリンクの構造に基づいて、原文書の分析を目的とした分類を行なった。

この分類ではリンクは大きくOrganizationalリンクとReferentialリンクに分かれる。

- Organizational 文書の構造を表現するリンク
- Referential ノード間にある、文書の構造に関するもの以外の関係

さらに、Organizationalリンクは以下のように分かれる。

- Aggregational 部分-全体関係を表現し、複合ノードを定義するリンク
- Grouping 複数のノードの集まりであるノードを定義するリンク

リンクが表現する関係の例を表1に示す。

表1: リンクが表現する関係

リンク名	リンクが表現する関係の例
Aggregational	「辞典の各エントリは、見出しと語義から成る」という関係
Grouping	「類語辞典の上位カテゴリは複数の下位カテゴリに詳細化されている」という関係
Referential	「右手」という単語と「左手」という単語の間にある[反対語]という関係

3 データダイアグラム

分析時に用いるダイアグラムには、原データ中にimplicitに表現されているリンクをexplicitに表現できる表現能力が要求される。またそのリンクも、タイプがあたり様々な属性を持っているのが普通であり、そのようなリンクの内部構造も表現できなくてはならない。

当初我々は、ERダイアグラム<sup>[5][6]</sup>を用いて原文書のモデル化を行なっていた。しかし、ERダイアグラムには以下の欠点があった。

1. 複合ノードの表現ができない
2. is-a関係が表現しにくい
3. リンクのタイプを明確に表現できない

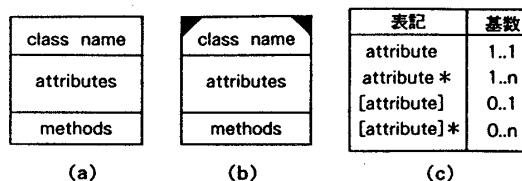
これに対して、Coadらのダイアグラム<sup>[1]</sup>は、Assembly Structure, Classification Structureによって上記の1., 2.を解決している。そこでこれをハイパーテキストの立場から整理し、2章で述べたリンクの分類に基づいて拡張を行なったダイアグラムを用い、原文書のモデル化を行なうことにした。以下我々が用いたダイアグラムの表記法について述べる。

3.1 クラス

クラスは以下の3つの要素によって構成される。

- class name そのクラス固有の名前
- attributes そのクラスのインスタンスが持つデータ
- methods そのクラスに対して指示できる操作

クラスの表記法は、Coadらのオブジェクト表記法<sup>[1]</sup>とほぼ同一である(図1(a))。ただし、抽象クラスを図1(b)、各attributeの多値性を図1(c)のようにして表現する拡張を行なった。



(a) (b) (c)

図1: クラスの表記法

\*The HyperDictionary Hydra(2) — Data Model —

†Jiro Nemoto, Shinji Yamashita, Yoichi Matsuyama, Hiroshi Naito, Koichi Masegi

‡Information Systems Research Center, CANON Inc.

また、クラス間の継承関係を表すis\_aリンクは図2のようにして記述する。

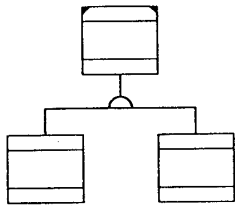


図 2: is\_aリンクの表記法

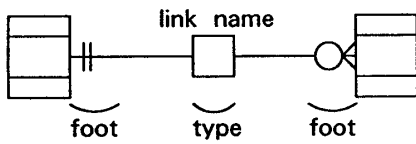
### 3.2 リンク

標準的なリンクは以下の3つの要素によって表現される。

- link name そのリンク固有の名前
- type そのリンクのタイプ
- foot 写像基数

2章で述べたリンクの分類がリンクの構造に基づいたものであったのに対して、link nameはその意味を反映したものを付ける。

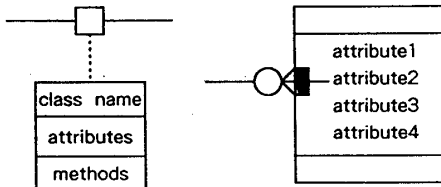
リンクの表記の例を図3(a)に、type及びfootの表記法を図3(b)に示す。リンクにもいくつかのattributesが付くが、デフォルトで与えられるattributesはダイアグラム上に記述しない。しかし、特別なattributesやmethodsが付け加えられたリンクは図3(c)のようにして記述する。また、embedded linkに関しては、どのattributeの中に埋め込まれているかを示すために、図3(d)のようにして記述する。



(a)

type	Aggregational	Grouping	Referential
	親 ◁ 子	集まり ⊕ 要素	◻
foot	⊥ (1..1)	⊥ (1..n)	⊥ (0..1)
			⊥ (0..n)

(b)



(c)

(d)

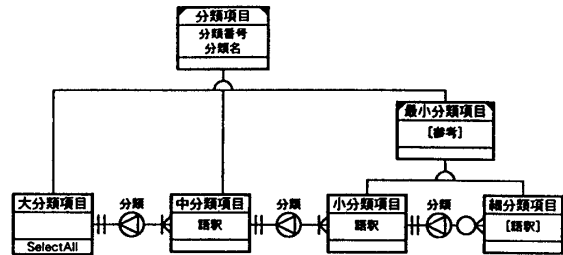
図 3: リンクの表記法

また、我々はリンクの記述に以下の制約を設けている。

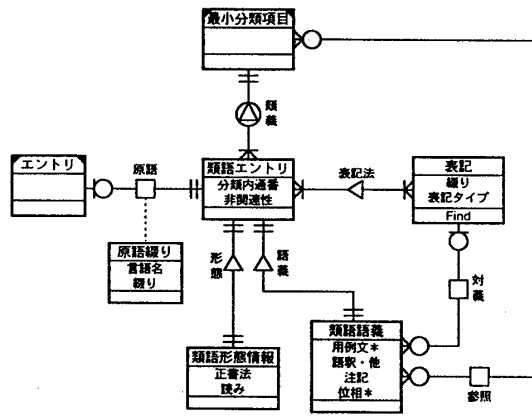
1. 一つの親クラスから複数のGroupingリンクが出ることはない
2. 一つの親クラスからAggregationalリンクとGroupingリンクが同時に起こることはない

### 4 モデル化の例

3章で述べたダイアグラムを用いて行なった類語辞典<sup>[7]</sup>のモデル化の例を図4に示す。この類語辞典は、属するカテゴリによって各エントリが分類されており、かつ、そのカテゴリ間には階層関係がある。そこで、カテゴリの階層関係を表すダイアグラム(図4(a))と、各エントリの構造を表すダイアグラム(図4(b))とによってモデルを表現している。



(a)



(b)

図 4: 類語辞典<sup>[7]</sup>のデータモデル

### 5 おわりに

辞書ハイパーテキストHydraの開発の際に採用したリンクの分類と、それに基づいた辞書のデータモデルを表現するための分析用ダイアグラムについて述べた。今後は辞書だけでなく、マニュアルや論文のような一般ドキュメントのモデル化についても検討していく予定である。

#### 参考文献

- [1] Coad,P.,Yourdon,E. "Object-Oriented Analysis", Prentice-Hall,1990.
- [2] DeRose,S.J. "Expanding the Notion of Links", Proceedings of Hypertext '89,ACM press,1989,pp.249-257.
- [3] Parsaye,K. et al. "INTELLIGENT DATABASES", John Wiley & Sons,Inc.,1989.
- [4] Conklin,J.,Begeman,M.L. "gIBIS:A Hypertext Tool for Exploratory Policy Discussion" ACM Transactions on Office Information Systems,Vol.6,No.4,Oct 1988,pp.303-331.
- [5] Martin,J.,McClure,C. "ソフトウェア構造化技法", 近代科学社,1988.
- [6] 酒井, "情報資源管理の技法", オーム社,1987.
- [7] 大野, 浜西, "類語国語辞典" 角川書店,1985.