

6H-3

# Continuous Object 概念に基づく ハイパーテキストデータベースモデル\*

銭 晴

神戸大学大学院自然科学研究科

田中 克己

神戸大学工学部

## 1 まえがき

ハイパーテキストの大規模化や実際の業務への本格的な適用の機会の増加につれて、従来のハイパーテキストの弱点 [3], 例えば、大規模ハイパーテキスト中での *disorientation*, 文書情報の追加・削除・更新に伴うリンクの修正・管理の複雑化, 集成的検索言語機能の欠如, マルチユーザアクセスに伴う並行制御機能の欠如などが指摘されている。

このため、最近、データベース技術とハイパーテキスト技術を統合してこれらの弱点を克服しようという、いわゆる、ハイパーテキストデータベースの研究が盛んになりつつある。例えば、種々の文書型やリンク型の管理に OODBMS を用いたり [7], 現在の位置やリンクを指示できる質問言語を開発したり [1], リンクを質問対で表現することで管理の手間の低減化を図る [5] などの試みや、超グラフを用いたデータモデルの提唱 [6, 8], ノードやリンクの抽象化機構の提案 [2, 4] などがある。

本稿では、2段階のスキーマ構成と一般化されたリンクの概念に基づく、大規模なハイパーテキストデータベースの構築・管理に適したデータモデルを提案する。本モデルは、OODBMS などによって実現された物理的なハイパーテキストデータベースのレベルの上に、リンク付けなどのオーサリングを行なう概念的なレベルを持つ2層構成であり、動的・多義的なリンクの設定や、リンクの端点となるオブジェクトの *identity* をデータベース作成時よりも後で決定できるなどの特徴を有している。

## 2 2段階スキーマ

2段階スキーマ構成は、型スキーマとビュースキーマからなる (Fig.1 参照)。型スキーマは通常の OODBMS が提供するクラス階層と、構造的リンク群からなる。構造的リンクとは、主にクラス定義時に決定されるオブジェクト間の Part-of 関連や、オブジェクト間をつなぐハイパーテキスト的なリンクを意味する。OODBMS が提供するクラス階層は文書の論理構造や属性構造の表現には適しており、型スキーマは主にこのような文書構造の表現や物理的に格納されたオブジェクト間のリンクの定義などから構成される。型スキーマのレベルで定義されているオブジェクトのことを、他レベルのオブジェクトと区別するために、特に、物理オブジェクトと呼ぶ。

ビュースキーマは、物理オブジェクト群から定義される、仮想的なオブジェクト群の分類のための種々の階層構造と、仮想的なオブジェクトをつなぐリンクの定義情報などから構成される。この、仮想的なオブジェクトを、概念オブジェクトと呼ぶ。概念オブジェクトの定義やグルーピングは、陽

に指定したり、また、質問などを用いて内包的に定義する。概念オブジェクトを少なくとも始点か終点にもつリンクのことを非構造的リンクと呼ぶ。

2段階スキーマを考える動機は、現状の C++ をベースとする OODBMS の次のような弱点である。(a) 文書の分類階層としては、クラス階層は制限が強すぎる (オブジェクトの所属クラスは1つに限定、同一クラスの全オブジェクトは同一のデータ構造, oid を保ちながらのオブジェクト移動が困難など), (b) OODBMS のリンク構造はクラス定義時に決定され、柔軟性に欠ける, (c) 任意のオブジェクトの仮想的なグループ化, すなわち, *view* の機能が欠如しており、文書の分類などには非力である。

概念オブジェクトは、すでに OODB 上でその構造が定義されている物理オブジェクトとは異なり、物理オブジェクト (例えば Paragraph 型オブジェクトの1部であったり、また、複数の物理オブジェクト (例えば、複数の Word 型オブジェクト) にまたがった形で定義されるものである。この意味で、文書の有する論理構造とは独立に、文書全体を一連の *continuous* なオブジェクトとみなして、その任意の部分をオブジェクトとして認識したものである。従って、概念オブジェクトに対応するテキストは重なりや包含関係があってもよい。概念オブジェクトを陽に定義する場合には、

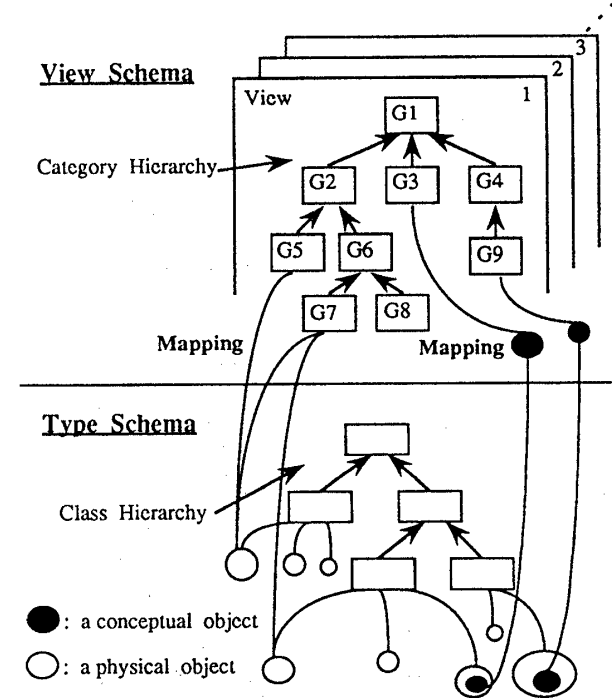


Fig. 1 Two Level Schemata

\*A Hypertext Database Model Based on Continuous Objects by Qing Qian and Katsumi Tanaka, Kobe University.

該当する物理オブジェクト (例えば, 1つの Paragraph オブジェクト) と, その中での位置を陽に指定すればよい. 概念オブジェクトの identity は, ユーザがクラス定義時よりも後に定義・制御できる必要がある. 例えば, 次の例は,

```
DEFINE Conceptual Object 'OODBMS'
FROM AcademicReport
WHERE 1984 < Year
```

1984 年以降に出版された AcademicReport 型文書オブジェクトに出現するあらゆる 'OODBMS' という単語を, 同一の概念オブジェクトとして定義するものである.

### 3 リンク

リンクは概念オブジェクトや物理オブジェクトの oid 集合の順序対  $\lambda = \langle \sigma_s, \sigma_t \rangle$  とし, 前者をソース, 後者はターゲットと呼ぶ. ソースやターゲットは外延的に (ユーザが陽に), または, 内包的に (質問として) 定義される. リンクの意味は, ソース集合とターゲット集合を各々節点集合とみなしたとき, これら2つの節点集合からなる完全 bipartite グラフで与えられる (Fig.2(a)). オブジェクト  $o$  のリンク  $\lambda$  を通じてのナビゲーションは,  $\lambda(o)$  と表し,

$$\lambda(o) = \begin{cases} \sigma_t & o \in \sigma_s \\ \text{undefined} & \text{otherwise} \end{cases}$$

と定義する. さらにオブジェクト集合  $O$  に対しても

$$\lambda(O) = \begin{cases} \sigma_t & O \subseteq \sigma_s \\ \text{undefined} & \text{otherwise} \end{cases}$$

とする (Fig.2(b)). これにより, リンクを用いて他のリンクを定義することができる. 例えば, Fig.2(c). さらに, リンクの合成 (Fig.2(d)) など, リンクに対する種々の演算が定義可能である.

リンクは, ソースまたはターゲットに現われるあらゆるオブジェクトが物理オブジェクトの場合, 構造的リンクといい, それ以外の場合, 非構造的リンクという.

リンクのソースおよびターゲットのオブジェクト集合が外延的に指定されている場合, このリンクを静的リンクといい, それ以外の場合, 動的リンクという. 動的リンクのソースやターゲットは, データベースへの質問という形で実現できる [5].

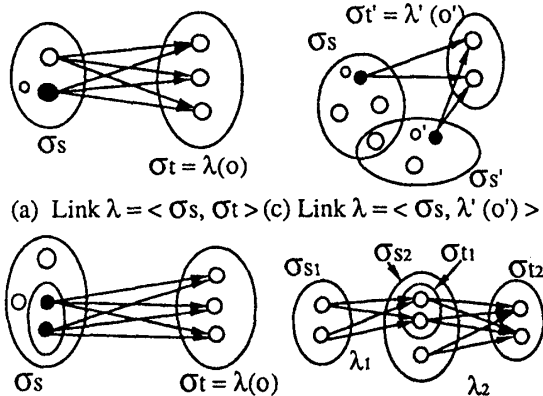


Fig. 2 Links in Hypertext Databases

リンクはさらに, 具体化リンクと多義的リンクに分類できる. 多義的リンクは, 直観的には, 複数の動的リンクのためのテンプレートであり, より具体的には, 与えられたオブジェクトから1つの動的リンクへの部分関数として定義される. 多義的リンク  $\tilde{\lambda}$  は次のように, 共通の変数  $x$  をもつ質問対として実現できる.

$$\tilde{\lambda} = \langle q_s[x], q_t[x] \rangle$$

多義的リンクのオブジェクト  $o$  による具体化とは, 変数  $x$  に  $o$  を代入して具体化された質問の対  $\tilde{\lambda}(o) = \langle q_s[o], q_t[o] \rangle$  と定義し, 具体化された多義的リンクは, 通常の動的リンクと同様の形で,  $(\tilde{\lambda}(o))(o')$  という形で利用できる.

例えば,  $q_s[x]$  を,  $x$  というフレーズを含む文書を検索する質問,  $q_t[x]$  を, その属性 KEYWORD の属性値として  $x$  をもつ文書を検索する質問とし, これらの対からなる多義的リンク  $\tilde{\lambda} = \langle q_s[x], q_t[x] \rangle$  は, Fig.3 と以下に示すように, ナビゲーションを開始するオブジェクトによって, その行き先を変えることができる.

$$(\tilde{\lambda}('OODB'))(o_1) = \{o_4, o_5\} \quad (\tilde{\lambda}('RDB'))(o_6) = \{o_8, o_9\}$$

### 4 あとがき

今後の課題として, (1) 一般化されたリンクの操作系とその表現能力, (2) 概念オブジェクトの定義のための機構, (3) 概念オブジェクトの object identity の決定機構などが考えられる.

### 参考文献

- [1] Fuller, M., et al., Proc. of DEXA'91, Aug. 1991.
- [2] Garg, P., CACM, Vol.31, No.7, July 1988.
- [3] Halasz, F., CACM, Vol.31, No.7, July 1988.
- [4] Hara, Y. et al., Proc. of DEXA'91, Aug. 1991.
- [5] Tanaka, K. et al., Proc. of the 7th IEEE Data Eng. Conf., April 1991.
- [6] Tompa, F.Wm., ACM TOIS, Vol.7, No.1, Jan.1989.
- [7] Wang, B. et al., Proc. of DEXA'91, Aug. 1991.
- [8] Watters, C. et al., ACM TOIS, Vol.8, No.2, April 1990.

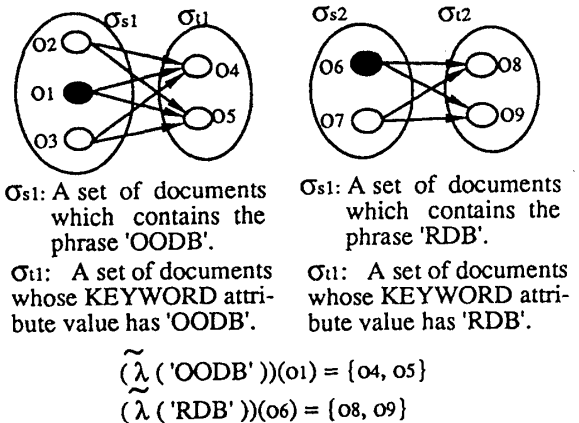


Fig. 3 A Polymorphic Link  $\tilde{\lambda}$