

ナレッジ・データベース構築新方式の一考察 - セル・データベースの適用 -

加藤 恵美子[†] 國井 利泰[‡]

要旨

ナレッジ・マネジメント・システム開発における基盤としてのナレッジ・データベースを、フレキシブルなセル・データベースで構築する方法を述べる。その前提としてのセル・データベース理論と、ナレッジ・データベースへの適用法を解説する。

On Novel Knowledge Database Architecture - An Application of Cellular DBMS -

Emiko Kato[†] and Toshiyasu L. Kunii[‡]

Abstract

We propose a novel architecture of knowledge databases as the foundation of knowledge management based on a flexible cellular database. A method to manage and apply knowledge databases is explained in the framework of the cellular database theory as the premise.

1. はじめに

現在、ネットワークに対して、ユーザの要請は1日24時間365日の連続稼働である。また、社内システムや固定会社間のような緊密なつながりを持つネットワークだけでなく、BtoB やジョイント企業間の連携などのゆるいつながりを持つネットワークへの対応の必要性を、ユーザが認識し始めている。

しかし、現状ではインターネットを通じたweb上のシステムは、プログラムで操作しているのみであり、データのアーキテクチャとしては、リレーショナル・データベースを初めとして、サイバーワールド上では展開できていない。

これらの問題点を解決する手法として提案されているのが、同値関係を基礎とするセル理論に基づいたセル・データベースである。これは従来の緊密なネットワークにも、また、新しく出てきたゆるいネットワークにも対応できるものである。具体的には、各属性の独立性が高まり、どんな状況にもあわせうる柔軟性を有している。

ここでは、まず第一に、このセル・データベースの理論を、モデル表現とDBMSの観点から述べる。

第二に、ナレッジ・データベースをセル・データベースで構築する方法を述べる。ナレッジ・データベースは諸々の変更柔軟に対処するために、生データで持つことが理想である。さらに知識は蓄積であり、かつ陳腐化したものは除いて集約する必要がある。この観点から、セル・データベースの履歴データと企業ポータル・システムを組み合わせることにより、ナレッジ・データベースを構築する。

[†] 法政大学大学院工学研究科電気工学専攻ITプロフェッショナルコース

IT Professional course, Graduate School, Hosei University

[‡] 法政大学大学院情報科学研究科

Graduate School of Computer and Information Sciences, Hosei University

2. セル・データベースのモデル表現

モデルとして、Adjunction Space Model とセルモデルを適用する。

第一ステップとして、Adjunction Space Model により、システム全体と同値関係を使用したマッピング関数を提示する。

第二ステップとして、セルモデルにより、属性を提示する。

モデル表現として、概念モデル、論理モデル、物理モデルのそれぞれの各段階に分けて説明する。

2. 1 概念モデル

I. 概念モデル作成にあたって前提となるシステム概念を、The Adjunction Space Model で表わす。

— The Adjunction Space Model —

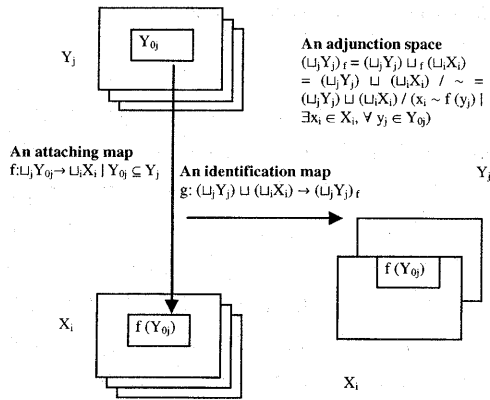


図1 概念モデル—The Adjunction Space Model

— Y_{0j} は構築したいシステムに必要なすべての属性の集まりを表わす。

— f, g は同値関係であることを表わす関数である。

—システム構築の対象となる形態が異なっても、最高レベルの概念は同一である。

ただし、 Y_j, X_i はそれぞれ以下を意味する。

(1) 企業（1社またはグループ）内システム

Y_j : 企業、 X_i : 部署

(2) BtoB

Y_j : 企業、 X_i : 企業

(3) BtoC

Y_j : 企業、 X_i : 消費者（個人または団体）

II. 概念モデルの詳細化を The Adjunction Space Model とセルモデルで表わす。

—セルモデル—

“ B^n_{kigyo} ” の形式で表わす。n は属性の数であり、システムなどを示す名前を付加できる。

(1) 企業（1社またはグループ）システム

—包含関係—

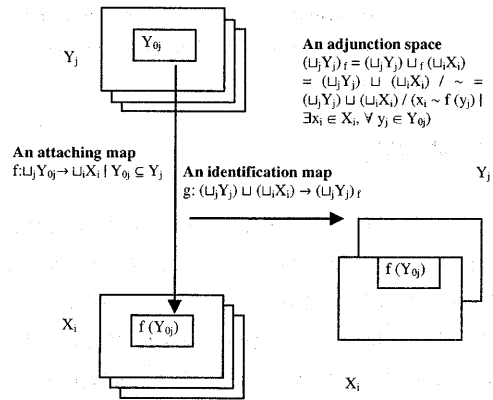


図2 企業システムの Adjunction Space Model

Y_j : 企業 = $B^n_{kigyo j}$

n: j の値ごとに任意の値を取りうる。

X_i : 部署 = $B^m_{bu i}$

m: i の値ごとに任意の値を取りうる。

Y_{0j} : 各システム（経理、販売など）= B^1_{system}

1: 各システムごとに任意の値を取りうる。

(2) BtoB

— 対等関係 (双方向) —

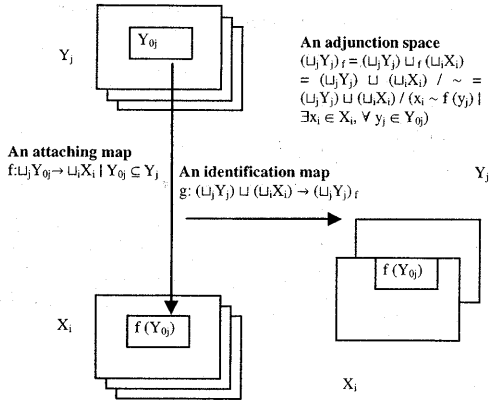


図3 BtoBのAdjunction Space Model

Y_j : 企業 = $B^n_{kigyoY_j}$

n : j の値ごとに任意の値を取りうる。

X_i : 企業 = $B^m_{kigyoX_i}$

m : i の値ごとに任意の値を取りうる。

Y_0 : 各システム (受発注など) = B^1_{system}

1 : 各システムごとに任意の値を取りうる。

— ただし、実際のシステム構築にあたっては下記のような状況に応じて詳細化して表わされることがある。

1) 売買取引などで企業内システムと連動する場合

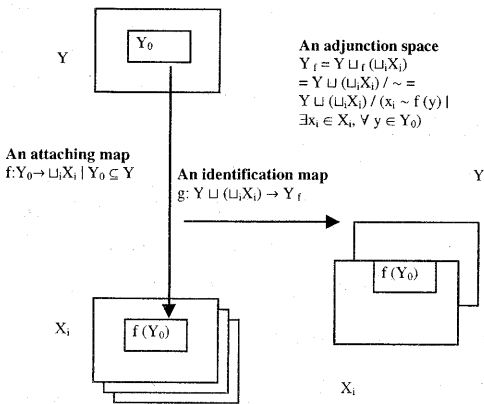


図4 BtoBで売買取引などで企業内システムと連動する場合のAdjunction Space Model

Y : 企業 = B^n_{kigyoY}

1 企業を表わす。

つまり企業ごとにこのモデルがある。

X_i : 企業 = $B^m_{kigyoX_i}$

m : i の値ごとに任意の値を取りうる。

Y_0 : 各システム (受発注など) = B^1_{system}

1 : 各システムごとに任意の値を取りうる。

2) 共同調達などで各企業内システムと連動する場合

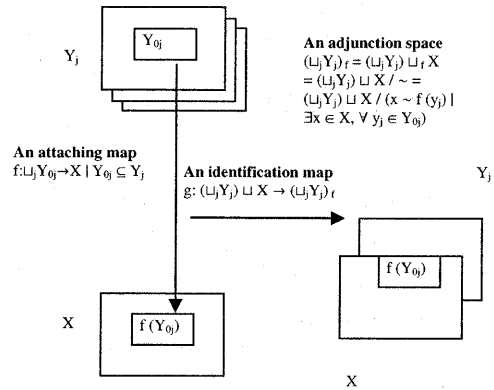


図5 BtoBで共同調達などで各企業内システムと連動する場合のAdjunction Space Model

Y_j : 企業 = $B^n_{kigyoY_j}$

n : j の値ごとに任意の値を取りうる。

X : 企業 = B^m_{kigyoX}

1 企業を表わす。

この企業のモデルも X, Y を置き換えた形で存在する。

Y_0 : 各システム (受発注など) = B^1_{system}

1 : 各システムごとに任意の値を取りうる。

(3) BtoC

—要求関係 (一方向) —

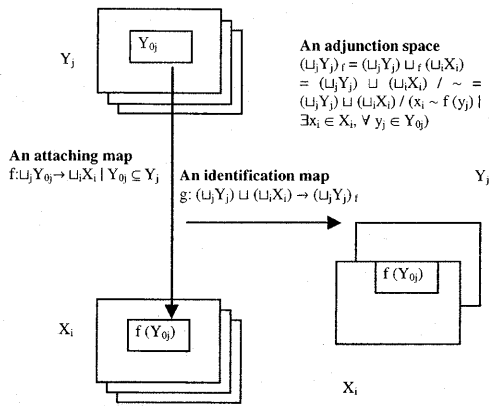


図6 BtoC の Adjunction Space Model

- Y_j : 企業 = $B^{n_{\text{kigyo}Y_j}}$
 n : j の値ごとに任意の値を取りうる。
- X_i : 消費者 = $B^{m_{\text{consumer}i}}$
 m : i の値ごとに任意の値を取りうる。
- Y_0 : 各システム (購買など) = B^1_{system}
 1 : 各システムごとに任意の値を取りうる。

—ただし、実際のシステム構築にあたっては下記のような状況に応じて詳細化して表わされることがある。

1) 企業内システムと連動する場合 (1企業)

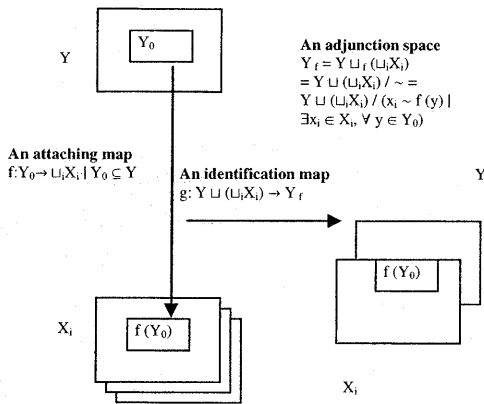


図7 BtoCで企業内システムと連動する場合 (1企業) の Adjunction Space Model

2) 各企業内システムと連動する場合 (複数企業)

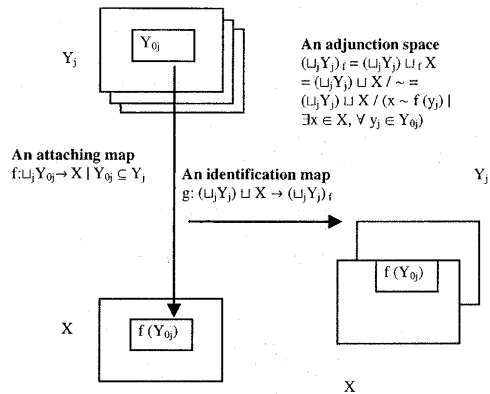


図8 BtoCで各企業内システムと連動する場合 (複数企業) の Adjunction Space Model

2. 2 論理モデル

The Adjunction Space Model とセルモデルで表わされた概念モデルから、モデルの詳細化を行なう際の手法の一例を述べる。

- (1) 属性
 - (2) 関数
 - (3) イベント (データ・フロー)
- 以上の3つの分析を行なう。

(1) 属性

セル・データベースでは、各属性の独立性が高いため順番などは関係なく、必要な属性をすべて定義するだけで良いが、確認作業などの便宜のためグルーピングを行なう手法が考えられる。これは、セルモデルのセル分割と同義である。エンドユーザには「属性グループ名+属性名」で管理するようツールで対応すれば良いと考えられる。

—セル分割—

$$B^1 = B^s + B^t$$

この時、 $1 = s + t$ である。

例)

B_{system}^1 システムに必要な全属性
 B_{group1}^s 属性グループ 1
 $B_{system1}^1(&B_{group1}^s)$ 属性 1 — 属性名 A
 $B_{system2}^1(&B_{group2}^s)$ 属性 2 — 属性名 B
:
 B_{group2}^1 属性グループ 2
 $B_{system3}^1(&B_{group2}^s)$ 属性 3 — 属性名 C
:
 $B_{system1}^1(&B_{group2}^s)$ 属性 1 — 属性名 Z

(2) 関数

同値関係を使用したマッピング関数はディクショナリに保管される。これにより、例えば異なったシステム間のデータの関連を管理できる。この管理もツールで行なう。

(3) イベント (データ・フロー)

プロセスの詳細化としては、イベント・ドリブンの手法を行なう。階層化も可能である。システム要件として関数で扱われる属性と、属性データの遷移のみ着目する。

2. 3 物理モデル

現行では既存のシステム資産を活かすためにも、セル・データベースのDBMSをリレーショナル・データベースへのフロント・エンド・プロセッサのように開発することが考えられている。その場合には、パフォーマンスや容量などの考慮として、例えばRDBのER図、テーブル設計(非正規化などを含む)などが必要となる可能性がある。

3. セル・データベースのDBMS

3. 1 実装について

- (1) 関数、属性間の対応などを管理するディクショナリを持つ必要がある。これにより、複合属性と単一属性との対応など、システムやデータベースの

変更をせずに、または最小限の工数で、異なるシステム間でもデータ共有が可能になる。各セル・データベースのDBMSが保持、管理をする。

- (2) データをOID(オブジェクトID)でのシステム管理を行なうことにより、削除時の履歴などにも有効なものとなる。
- (3) ホモトピー同値を保証するための変更履歴データの蓄積(OIDとカウンタでの管理)を行なう。つまり、データへの作成・更新・削除アクションの各1件ごとに履歴データに追加される。これにより、差異データをすべて持つことになるので、どの時点でもその時のデータの状態に戻すことが可能になる。そして、関連した他のシステムのデータ変更も容易に追跡可能になる。

3. 2 課題

- (1) RDBMSのフロント・エンド・プロセッサのように開発するような時には、非正規化フィールド多対多の関係などへの対処が必要になる。
- (2) 処理やデータ量の増加への対処(パフォーマンスやデータ容量の計算)が必要である。

4. ナレッジ・データベース

企業ポータル・システムのデータベースには、仕事に必要な文書、URL、検索結果などが、部署やチームや個人ごとに管理されている。そして、普通これらの内容は常に更新されている。しかし、例えばマーケティングやリサーチ部署と他部署とのように、部署間のタイムラグが発生する場合などでは情報が効果的に移行していない可能性がある。つまり、全社的に見れば重複した作業時間を費やしているかもしれない。有用な情報を逃がしているかもしれない状況である。

全社的に横断した情報を知識として活用するために、履歴として保持されている共有可能な情報を使用して行なおうというのが、このナレッジ・データベースである。つまり、セル・データベースの履歴データと企業ポータル・システムを組み合わせることにより、ナレッジ・デ

データベースを構築する。

このナレッジ・データベースを使うことによって、全社的に、社内情報、webを通した外部情報を一括管理できるナレッジ・マネージメント・システムの開発が可能となる。

エンドユーザ用と管理用の2つの機能を持つナレッジ・マネージメント・システムのエンドユーザ用システムを Adjunction Space Model とセルモデルで表わす。

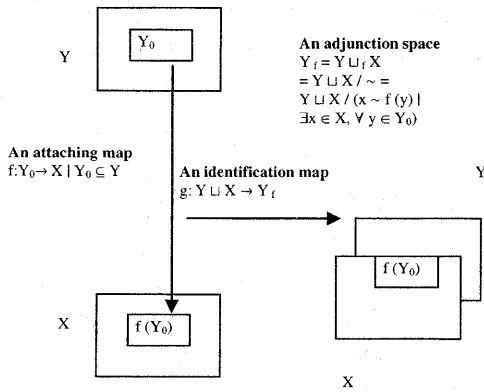


図9 ナレッジ・マネージメント・システムの Adjunction Space Model

このナレッジ・データベースは、

- 1) セル・データベースの DBMS により作成される、企業ポータル・データベースの履歴データベースである。
- 2) 物理的には複数のサーバに分散されている可能性もあるが、意味的には1つのものであるので1つとみなす。

Y が保持している知識と、X が得たい知識が同値の時、このシステムが成り立つ。

Y : ナレッジ・データベース $= B_{\text{kmDB}}^n$

X : 社員 $= B_{\text{syain}}^m$

Y₀ : 知識 $= B_{\text{itiki}}^1$

5. まとめ

以上のように、セル・データベースを適用して、ナレッジ・データベースを生データとして知識を蓄積し、かつ陳腐化したものは除いて集約することが可能である。このナレッジ・データベースから効果的に知識を取得するシステムの開発もまた可能である。

今回は、全社システムを対象に考察したが、セル・データベースを適用することで、将来的には他のネットワークと知識の共有をすることも考えられる。

6. 参考文献

- [1] Toshiyasu L. Kunii, "Web Information Modeling: The Adjunction Space Model", Proceedings of the 2nd International Workshop on Databases in Networked Information Systems (DNIS 2002), in press, The University of Aizu, Japan, December 16-18, 2002, Lecture Notes in Computer Science, Subhash Bhalla, Ed., Springer-Verlag, December, 2002.
- [2] Toshiyasu L. Kunii: A Memo on a Cellular Model for Cyberworlds Design (1999-2002).