

データベース操作の高度化に対する意味情報の表現

1Q-2

上原祐介 渡辺豊英 吉田雄二 福村晃夫  
 (名古屋大学 工学部) (中京大学)

1. はじめに

データ・モデルは、データベースの論理構造の表現モデルであるが、意味情報の表現モデルではない。従って、データ・モデルに対して定義された操作コマンドによってデータベースを処理する場合、利用者がデータベースの意味に関する知識(以後、意味情報と呼ぶ)を持っていないければ、効率的に操作することはできない。この問題に対して、自然言語を用いて、データベースの操作を高度化する自然言語インタフェース、または知的インタフェースの課題が研究された。しかし、この研究の主題は、自然言語処理に重点が置れた。本研究では、データベースの意味情報としてどのような情報を想定し、この意味情報をどのような表現構造として管理、処理するかという課題の下に、データベースの意味表現モデルを用いて、データベース操作の高度化を実現することを目的としている。本稿では、データベースの意味情報をデータ・モデルとして陽に表現するために構成されたデータベースの意味表現モデルを報告する。

2. 意味表現のための枠組み

データベースの表現は、そのデータベースに対応する実世界の「実体」、「実体集合」、及び「実体集合間の関連」として捉えることができる。従って、この「実体」、「実体集合」、「実体集合間の関連」の概念を用いてデータベースの意味情報を表現する。このデータベースの意味情報を表現する枠組みを意味表現モデルと呼ぶ。以下、この意味表現モデルを説明する。

意味表現モデルの定義

意味表現モデルはオブジェクト(object)、クラス(class)、リレーション(relation)により構成される。それぞれの構成要素を先に定義する。

・オブジェクト

実世界における実体をオブジェクトとする。オブジェクトは図1のように分類される。

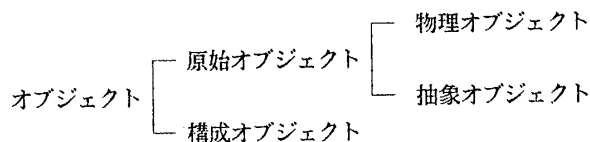


図1 オブジェクトのカテゴリ

オブジェクトは、原始オブジェクト(PrO)と構成オブジェクト(CoO)に分類される。PrOは、最小単位のオブジェクトであり、他のオブジェクトから構成されない。CoOは複数のPrOに集約操作を施して構成されたオブジェクトである。さらに、PrOは物理オブジェクト(PhO)と抽象オブジェクト(AbO)に分類される。PhOは実世界に実在する活動実体であり、「ある人」、「ある部品」などがこれに分類される。また、AbOは実世界に存在しない抽象的な概念実体であり、「ある人の名前」、「ある部品の番号」などがこれに分類される。

・クラス

判別可能な共通の特徴を有するオブジェクトの集合がクラスである。クラスはその特徴により意味付けされ、オブジェクトのカテゴリに従い、図2のように分類される。

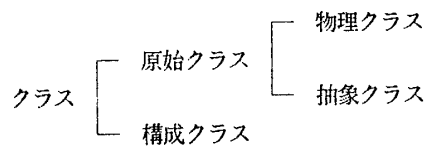


図2 クラスのカテゴリ

クラスは原始クラス(PrC)と構成クラス(CoC)に分類される。PrCはPrOの集合であり、CoCはCoOの集合である。さらに、PrCは物理クラス(PhC)と抽象クラス(AbC)とに分類される。PhCはPhOの集合であり、「供給者である」という共通の特徴を有したPhOの集合「供給者」などがこれに類する。AbCはAbOの集合であり、個々の部品の番号の集合「部品番号」などがこれに分類される。PhOには、これを指し示すためにAbOが付随する。このAbOのクラスを、ポインタ・クラスと呼ぶ。また、ポインタ・クラスにおいて、それが付随するPhCに対し、要素同志が1対1の対応関係にあるポインタ・クラスを特にキー・クラスと呼ぶ。PhCには少なくとも一つのキー・クラスが付随する。

・リレーション

クラス間の関連が、リレーションである。リレーションによって関係付けられるクラスには、そのリレーションにおける意味付けを表す役割(role)が割り当てられる。役割は格フレームに従って、図3に示した意味付けがある。リレーションは、

リレーション名、関連するクラス、及びそれらのクラスに割り当てられた役割により意味付けられる。リレーション、及びそのリレーションが関連付けるクラスに対して集約操作を施すことにより、構成クラスが生成される。このクラスの要素である構成オブジェクトは、リレーションで関連付けられるクラスから対を成して抽出され、タプルを生成する。

ロール名(略号)	意味
行為主 (agt)	行為を行うもの
被行為主 (recp)	行為を受けるもの
対象 (obj)	行為の対象
手段 (inst)	行為の手段
場所 (at)	行為の場所
時間 (time)	行為の時間
始点 (from)	行為の開始点
終点 (to)	行為の終了点

図3 役割とその意味

意味表現モデルの具体例として、次の内容を表すデータベースを考える。

「供給者は、ある一定量の部品を供給している。供給者は番号によって識別され、また名前を有している。部品は番号によって識別され、また名前を有している。部品には色がある。色は名前によって識別される。」

この内容を表す意味表現モデルは、次のようなクラス、リレーションにより構成される。

- ・物理的に実在する活動実体の集合である供給者、部品、色を表す物理クラス「供給者」、「部品」、「色」。
- ・抽象的な概念実体の集合である番号、名前、供給量を表す抽象クラス「番号」、「名前」、「供給量」。なお、クラス「番号」は、クラス「供給者」、「部品」のキー・クラスである。また、クラス「名前」はクラス「色」のキー・クラスであり、同時にクラス「供給者」、「部品」のポインタ・クラスである。
- ・クラス「供給者」とクラス「部品」間の関連を表すリレーション「供給」。なお、このリレーションに対して、クラス「供給者」には役割「行為主」、クラス「部品」には役割「対象」が割り当てられる。
- ・クラス「部品」とクラス「色」との間の関連を表すリレーション「所有」。なお、このリレーションに対して、クラス「部品」には役割「行為主」、クラス「色」には、役割「対象」が割り当てられる。
- ・リレーション「供給」、及びクラス「供給者」、「部品」によって構成される構成クラス「供給」。
- ・クラス「供給」とクラス「供給量」間の関連を表すリレーション「所有」。なお、このリレーションに対して、クラス「供給」には役割「行為主」、クラス「供給量」には、

役割「対象」が割り当てられる。

3. 意味表現のためのモデル図

意味表現モデルの表記法を以下に述べる。

- ・物理クラスは付随するポインタ・クラスとともに、上下二段に分割された四角形で表わされる。クラス名を下段に、ポインタ・クラス名を上段にそれぞれ記す。キー・クラスは二重枠により識別される。
- ・抽象クラスは四角形で表され、クラス名はその内部に記される。
- ・リレーションは菱形で表され、リレーション名はその内部に記される。また、リレーションによって関連付けられるクラスとそのリレーションは実線で結ばれ、そのクラスの役割名がその実線に付される。
- ・構成クラスは、これを構成するクラス、リレーションを囲む四角形で表される。クラス名はその枠外に記される。

第2節で例示した意味表現モデルのモデル図を図4に示す。

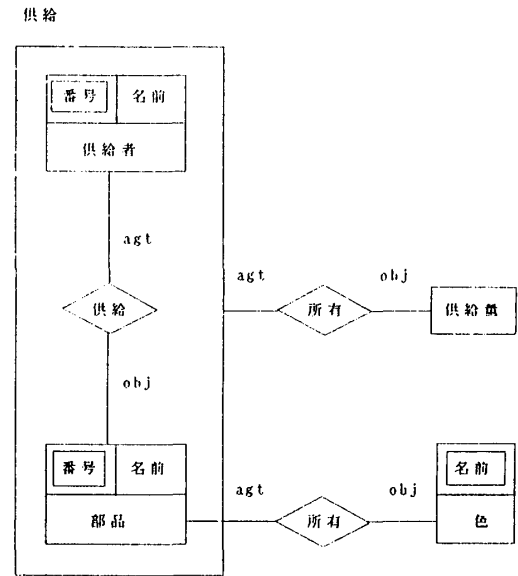


図4 意味表現モデルの表記例

4. おわりに

本稿では、データベース操作の高度化を目的としたデータベースの意味表現モデルを提案した。このモデルでは、データベースをデータ・アイテムと、それらの間の関連として捉えている。しかし、実際のデータベースでは、例えば「一貫性の制約」のように、実体、実体集合、関連だけでは捉えられない情報が存在する。本意味表現モデルにおいて、このような情報をどのように反映させるかを検討することが必要であり、今後の検討課題の一つである。

参考文献

(1) Ullman, J. D. :Principles of Database Systems, Second Edition, Computer Science Press, 1982.