

解説



データベース関連技術の標準化

6. 概念スキーマモデル機能 CSMF の標準化†

佐藤 英 人††

1. はじめに

本稿では、1994年にスタートしたISO（国際標準化機構）のCSMF（Conceptual Schema Modelling Facility：概念スキーマモデル機能）標準化プロジェクトを中心に、その背景、考え方、現状を解説する。

2. 概念スキーマモデル機能とは

「概念スキーマ」は、データベースの対象世界の論理的な構造を表現するものとして、ANSI/SPARCによって1975年に提案された概念¹⁾であり、今日では、データベース分野に限らず、情報システム分野全般で使われている用語である。「概念スキーマモデル機能」は、この概念スキーマを記述するための1組の概念の集まりを指している。

モデル機能（Modelling Facility）は、概念を記述する概念であるため、言葉で説明しようとするとき混乱をきたしやすい。そこで、RMDM（データ管理参照モデル²⁾）のレベル対を用いて、例示的にその位置付けを示したものが、図-1である。

現実世界の個（この例では、山田太郎と呼ばれる個人）と1対1に対応する記述が、レベル1のインスタンスレベルであり、それらを類型化したもの（この例では、従業員）が、レベル2のスキーマレベルである。このスキーマレベルにも似たようなものがたくさん登場してくる。会社をモデル化する場合であれば、従業員、商品、注文、倉庫などがでてくる。これらを類型化したものがレベル3のモデル機能レベルである。P. Chenの実

体関連モデル³⁾では、「実体」や「関連」が、このモデル機能レベルの概念であり、SQLでは、「表」や「列」がこのレベルの概念である。

従来、データベース分野では、このレベル3の概念の1組の集まりを単に「データモデル」と呼ぶことが多かった。しかし、レベル2のスキーマ概念の集まりも、「〇〇業務のデータモデル」と称することが、現実には多い。そこで、標準化に当たって、混乱を避けるために、レベル3の概念の集まりを「モデル機能」と呼ぶことにしたのである。

なお、「概念スキーマモデル機能」は、このモデル機能の中でも、SQLの表や列といった情報の表現を対象とするのではなく、実体や関連といった現実世界の対象に関する概念を扱うものである。従来、データベース分野で「意味データモデル」と呼ばれてきたものとはほぼ同じのものであると考えて差し支えない。また、オブジェクト指向モデリングで、概念モデルと呼んでいるものは、レベル2の概念の集まりであり、それらを記述するために用いられるオブジェクト、性質、状態、イベントといった概念は、レベル3の概念である。

3. 概念スキーマモデル機能標準化の目的と経緯

先に述べたANSI/SPARCによる3スキーマの提案以降、そこで規定された概念スキーマを記述するモデル機能として、多くの意味データモデルが提案された。P. Chenの実体関連モデルやG. M. NijssenのENALIM（今日のNIAM⁴⁾）などは、その代表的なものである。

1980年代のはじめにISOは、これらの概念スキーマのための概念・用語を標準化すべくプロジェクトを起こした。このプロジェクトは、1987年にTR 9007と呼ばれるテクニカルレポート⁵⁾

† Standardization of Conceptual Schema Modelling Facility (CSMF) by Hideto SATO (Faculty of Business Administration, Tokyo International University).

†† 東京国際大学商学部

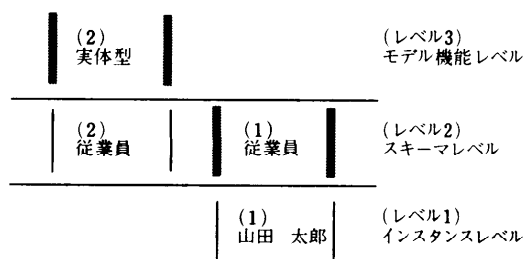


図-1 レベル対とモデル機能

を公刊して解散したため、TR 9007 プロジェクトと呼ばれている。

このレポートは、概念スキーマに係る問題を深く考察したもので、学術的価値の高いものであるが、結論的には3種概念スキーマモデル機能(実体関連モデル、2項関係モデル、述語論理モデル)を併記し、拘束力のないテクニカルレポートに止まったため、標準化という意味では実世界に大きな影響を与えることはなかった。

その後、1990年代に入って、別の意味で概念スキーマモデル機能の標準化が求められるようになってきた。それは以下のような情報システム分野の新しい動向に対応するものである。

(1) データ交換上の必要性

商取引の電子化である EDI (Electronic Data Interchange) の普及・拡大にともない、交換するデータの意味を記述する情報の標準化が必要となってきた。EDI は単にデータの送受信に関するプロトコルだけでなく、取引の形態や取扱い商品の性質に応じたビジネスプロトコルを必要とする。このため、交換されるデータの意味や交換後の取扱いを規定できる概念スキーマモデル機能が必要になる。

(2) CASE ツール統合上の必要性

ソフトウェア開発を自動化する CASE (Computer Aided Software Engineering) ツールの普及にともない、ツール間の連携をとり、定義情報を共有するための標準化が必要になってきた。CASE ツールは、概念レベルの上流ツールから実装レベルの下流ツールに至るまでの多くのツール群で構成される。各レベルのツールは、それぞれより上位のツールの出力を入力として受け取る。ツール相互間の互換性を高め、円滑な連携を可能とするためには、定義情報の意味を規定でき

る標準的な概念スキーマモデル機能が必要である。

(3) 概念モデリングの普及

業務ソフトウェア開発におけるデータベース利用の一般化、CASE ツールやリポジトリの登場、オブジェクト指向概念の普及にともない、ソフトウェア開発における概念レベルのモデル化は不可欠のものとなってきた。この結果、概念モデリングに関する種々の方法論が登場し、それぞれが独自の概念スキーマモデル機能を提案するに至っている。

(4) 国際標準間の重複

(1)~(3)のような背景から、文献6)で詳しく述べられているように、ISOをはじめとする標準化組織において、それぞれの分野で概念レベルを含むモデル機能の標準化が進行している。しかしそれらは、多くの共通点を持つにもかかわらず、それぞれの分野の対象に適した固有の概念・用語を用いており、標準そのものの統合や融合あるいは共存が困難になっている。

以上の状況の中で、ISO では、1989年11月のイタリア・フローレンス会議を皮切りに、概念スキーマモデル機能標準化の新プロジェクトの検討が始まり、1993年6月の横浜会議で新プロジェクトの提案が行われた。投票の結果、5カ国(日、米、カナダ、ベルギー、フィンランド)の賛成が得られ、1994年7月の英国サウザンプトン会議で SC 21 WG 3 の下の活動として、CSMF プロジェクトが発足した。

4. CSMF プロジェクト

4.1 標準化作業の現状

CSMF プロジェクトは、発足と同時に作業を開始し、現在、1996年5月に米国カンザスシティで開催されるラポータ会議でワーキングドラフトを完成させるべく、作業を進めているところである。

ラポータはカナダの Baba Piprani、エディタは米国の John K. Sharp と日本の穂鷹良介が務めている。常時参加している国は、カナダ、米国、日本、イギリス、オランダ、フランスの6カ国であり、他にスウェーデン、オーストラリア、韓国、ブラジルが1~2回参加している。

作業は、米国が中心となってワーキングドラフ

ト案を用意し、主として、日本、イギリス、オランダがこれに対し対案や修正案を提案するという形で、進められている。CSMFは対象が抽象的なものであるだけに、各国代表間の相互理解に時間がかかったが、ようやく大筋について合意が得られつつあるところである。

4.2 標準CSMFの位置付け

3章で述べたように、CSMFの応用領域と想定されているEDIやCASEなどの分野では、すでに多くの国際標準が制定されている。標準CSMFはこれらの既存の標準のモデル機能に取って代わることを意図するものではなく、これら相互間の関係を記述する手段を規定することにより、異なるモデル機能間のスキーマ情報の交換、さらには、インスタンスデータの交換を容易にすることを目的としている。

この関係を図示したものが図-2のアーキテクチャである。この図は現在審議中のワーキングドラフト⁷⁾に収録されているものであるが、まだ完全な合意が得られているものではない。今後字句の変更などが施される可能性は高いが、大筋は変わらないものと思われる。

この図で、最下層のApplication Modelと書かれたレベルには、「〇〇会社のスキーマ」「△△業務のスキーマ」といった個々の应用スキーマが位置付けられる。先の図-1でいえばレベル2のスキーマレベルに相当している。その上のMeta Modelと書かれたレベルは、应用スキーマの定義言語を規定するレベルで、図-1のレベル3のモデル機能レベルに対応している。最上位のMeta-meta ModelレベルにあるDefining Schemaは、個々のモデル機能を定義する言語の規定であり、自然言語で補完された一階述語論理のようなものが想定されている。

Defining Schemaが提供するものは、概念定義のための構文に過ぎない。そこで、概念の意味を規定するための最小限のプリミティブを集めたものがMeta Modelレベルに書かれているNormative Schemaである。標準CSMFはDefining Schemaで規定される構文(Syntax)とNormative Schemaで規定される意味定義(Semantics)で構成される。

個々の应用分野のモデル機能は、この図のMeta Model 1, Meta Model 2, …である。これ

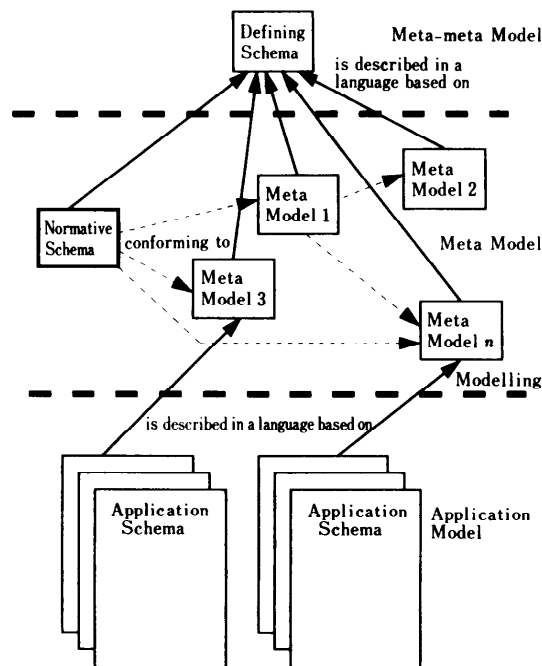


図-2 CSMFアーキテクチャ(暫定案)

らのモデル機能は、標準CSMFとは独立に、それぞれの分野の应用スキーマを規定できる。しかし、それぞれのモデル機能の構成要素をDefining Schemaの構文とNormative Schemaの要素を使って定義することにより、異なるモデル機能間でのスキーマ情報の同一性の判定や(可能な限りの)スキーマ変換ができるようになる。

4.3 標準CSMFの具体的内容

上で述べたように、標準CSMFは、モデル機能を定義する構文(Defining Schema)と基本的なモデル概念(Normative Schema)で構成される。以下、それぞれについて、これまでに行われた議論の概要を紹介する。

4.3.1 抽象概念スキーマ言語

モデル機能を定義する構文の規定には、述語論理の活用が有益であるという点については、各国とも異論がない。この面における具体的な提案としては、米国提案の「抽象概念スキーマ言語」(ACSL: Abstract Conceptual Schema Language)がある。これは、人工知能分野の知識交換を目的としてスタンフォード大学が中心となってまとめたKIF(Knowledge Interchange Format)⁸⁾をもとに、それを簡略化したものである。内容としては、述語論理をLISPの構文で書

けるようにしたものである。

LISP 構文であるため、慣れないと読みにくいものであるが、電子的情報交換が容易であるという利点がある。標準 CSMF にこの ACSL を採用するか、それとも、通常の述語論理の記法をそのまま用いるかは今後の議論の課題である。

4.3.2 規範的構成要素

構文だけ決めても概念を定義することはできない。これは文法があっても単語がなければ文を構成できないのと同じである。そこで CSMF プロジェクトでは、概念スキーマモデル機能を定義する上で必要最小限の基本的な概念の規定を進めている。この基本的な概念のことを規範的構成要素 (Normative Constructs) と呼んでいる。

このとき、何を規範的構成要素と見るかについて、これまでに3つの異なる意見がでている。

1) 概念の基本的意味要素への分解

CSMF プロジェクトが正式発足する前に、米国は J. F. Sowa の CG (Conceptual Graph: 概念グラフ)⁹⁾ を用いた標準 CSMF を強力的に提案した時期がある。この CG は、述語論理に基礎をおく意味ネットワーク型の概念モデル機能で、概念間の関係記述を C. J. Fillmore の格文法 (Case Grammar) に基づく約 40 の基本的な関係およびその合成に制限するところに特徴がある。

この CG を用いて定義された概念間では、意味的な共通性と差異を機械的に抽出できる。つまり、スキーマ変換の自動化が(ある程度)可能である。このアプローチは理論的に優れたものであるが、個々の概念の基本的な関係への分解は、人手に依らざるを得ないところがあり、実際のかどうか疑問が呈されている。

2) 既存モデル機能の概念の体系化

第2のアプローチは、既存のモデル機能で用いられている概念を網羅的に集め、それらを意味的に分類し、体系化しようとするものである。CG を取り下げた後の米国提案はこれに近いものであった。このアプローチは、公平ではあるが、規範的構成要素間の重複や規範的構成要素の矛盾した複数定義を生み、その規範性が疑われるものになりかねない。また、作業的にも、どこまでやれば十分か明確でないという難点がある。

3) モデル機能間の共通要素の抽出

世の中に異なるモデル機能がたくさんあるとい

っても、そこで用いられているモデル概念には見かけほど大きな差はない。モデル化の対象を表現するのに、実体、実体型、オブジェクト、クラスなど異なる用語が用いられていても、その意味に大きな差があるわけではなく、若干の機能的な差があるに過ぎない。属性、性質、関連、リンク、参照などのモノとモノとの結び付きについても同様である。

そこで、多くのモデル機能間で共通に見られる概念を抽出し、それを標準化すれば、他のモデル機能については、それとの対比で、同一性と差異を定義することができる。これが第3の立場であり、主として、日本が主張してきたものである。このアプローチは、1) ほど理論的ではないが、実際的である。簡単に対応関係が付けられないような概念相互の変換が実務上必要であるケースは稀であると考えられるからである。

CSMF プロジェクトの議論は、1) から始まり、2) を経て、現在3) に近いところまでまとまりつつある。昨年の暮れに開かれた中間会議で、一応の合意が得られた規範的構成要素の案は、図-3 に示すとおりである。

しかし、これらの要素はまだ単に列挙されたレベルに過ぎず、その具体的な仕様は今後の課題である。仕様の具体化にともなって、これらの要素が変更される可能性は高い。この現時点の規範的構成要素は、伝統的な ER (実体-関連) やオブジェクトの概念に基づくものではなく、ファクト

Normative Constructs の案 (SC21 WG3 N1909)

- 1) **Data concepts/Enterprise data**
entity-type, supertype/subtype, entity-instance, fact-type, role, attribute-type, attribute-instance, uniqueness-constraint, value-constraint, mandatory-constraint
- 2) **Process concepts/Enterprise activities**
process, event, trigger, state, agent, input/output
- 3) **Temporal consideration/Enterprise events**
relative-time, causal-dependency
- 4) **Business rules**
algorithm
- 5) **Other**
universe of discourse, conceptual schema

図-3 規範的構成要素 (暫定案)

(事実) をベースとする色彩が強いものである。

4.4 今後の予定

1995年7月のオタワ会議で勧告されたCSMF標準化の予定は、以下のとおりである。

WD (ワーキングドラフト)	96年1月
CD (委員会草案)	96年7月
DIS (国際標準草案)	97年7月
IS (国際標準)	98年7月

しかし、1996年3月時点で、まだWDが固まったといえる段階ではない。この予定どおりに標準化プロセスを進行させることは難しい状況にあるといえよう。

5. むすび

以上、現在進行中のISOのCSMFプロジェクトについて、その背景と現状を紹介した。CSMF標準は、他の情報関連の標準と異なり、ただちに具体的な製品に結び付くような性質のものではない。しかし、情報関連の種々の分野で概念レベルの標準化が進行しつつある現在、その早期の標準化が求められている。まず、必要最小限の部分に的を絞って、早期実現を目指すという日本の立場を今後も堅持して、プロジェクトに協力していきたいと考えている。

謝辞 末筆ながら、当プロジェクトに参加・協力いただいているSC 21/WG 3日本委員会の各委員、ならびに、情報処理学会情報規格調査会、日本規格協会の各委員に感謝の意を表します。

参考文献

- 1) ANSI/X 3/SPARC: Study Group on Data

Base Management Systems: Interim Report, FDT (Bulletin of ACM-SIGMOD), Vol. 7, No. 2, (1975).

- 2) ISO/IEC 10032: Reference Model of Data Management (1995).
- 3) Chen, P. P.: The Entity-Relationship Model—Toward a Unified View of Data, ACM Transactions on Database Systems, Vol. 1, No. 1 (1976).
- 4) Nijssen, G. M. (ed.): Architecture and Models in Data Base Management Systems, North-Holland (1977).
- 5) ISO/TR 9007: Information Processing Systems—Concepts and Terminology for the Conceptual Schema and the Information Base (1987).
- 6) 穂鷹良介: データベース関連技術の標準化の概要, 情報処理, Vol. 37, No. 7 (1996).
- 7) ISO/SC 21 WG 3 N 1909: Information Technology—Conceptual Schema Modelling Facilities (1996).
- 8) Genesereth, M. R. and Fikes, R. E.: Knowledge Interchange Format Version 3.0 Reference Manual, Stanford University (1992).
- 9) Sowa, J. F.: Conceptual Structures, Addison-Wesley (1984).

(平成8年4月4日受付)



佐藤 英人 (正会員)

1947年生。1970年東京大学工学部都市工学科卒業。経済企画庁に入庁。1978年筑波大学社会工学系専任講師。1984年大阪大学社会経済研究所助教授。1989年東京国際大学商学部経営情報学科教授。工学博士(東京大学)。データベース、オブジェクト指向モデリングの研究に従事。ISO SC 21/WG 3/CSMF SG 主査。日本規格協会情報資源スキーマ調査研究委員会委員。著書「統計データベースの設計と開発—データモデルと知識ベースの応用」。ACM 会員。