

概念辞書とEDI

玉川義人
筑波大学

実体素性・役割を階層的に表現した概念辞書を考えることで異なる業界で考えられているEDIデータ項目の役割としての意味的な共有化が図られる。これにより、業種・業界に寄らないレベルでの概念辞書の有効性が示される。また意味的制約条件の共有化に対しての有効性も示される。

情報システム設計で現れる実体型を抽象化し、実体素性としてとらえ、その階層構造を与えることでデータの部品化・共有化を概念辞書は目指している。これはオブジェクト指向的なデータ中心のシステム設計と目指すところは同じである。

THE CONCEPT DICTIONARY AND EDI

Yoshito Tamagawa
University of Tsukuba
Tennodai, 1-1-1, Tsukuba-shi, Ibaraki 305, Japan

We can attempt to share EDI data items in different business fields semantically by making a concept dictionary that represents the entity primitive and the role hierarchically. Therefore the advantage of the concept dictionary is shown in terms of the data level independent of business fields or areas. And also, it is shown in terms of the sharing of semantic constraints.

We will aim for the partitioning and the sharing of data by the concept dictionary so that we have to clear up the entity primitive by abstracting the entity type appearing in the design of the information system and show the hierarchy.

This prospect is the same as that of data oriented system design methods which are object oriented.

1. はじめに

シソーラスとは、語句を意味によって分類排列し、各語句について同意語、類義語、上位語、下位語、反義語、対義語などを記述した辞書である。情報検索においては、語の意味関係に注目して整理した統制語を集めた検索用語集として、シソーラスは重要なものとなっている。

先に、筆者らは実体素性及び役割の階層性に注目して概念を整理したものを、概念辞書とした¹⁾。これも、シソーラスと考えられるものである。本稿では、こうした概念整理を図ることのEDIに対しての有効性について述べる。

2. EDIビジネスプロトコール

近年の情報ネットワークの進展につれて、端末から通信回線を介してデータを伝送して取引を行う電子取引が活発となってきている。これがすなわち、EDI (Electronic Data Interchange) である。そして、EDIは業務の効率化・生産性の向上に供することが期待されている。

そのため、各業界においてEDI構想が推進されている^{2),3),4)}。EDIによるメリットとしては、建設産業情報ネットワーク推進協議会が構築するCI-NETにおいて、次の事柄が挙げられている⁴⁾。

- a.事務作業の効率化
- b.人為ミスの削減
- c.連絡時間の短縮
- d.企業経営の効率化

aは、情報が直接社内システムに送られてくるため、帳票から社内システムに対しての再入力が必要なくなるということである。bは、ペーパーレス化、すなわち、電子媒体を使うことにより転記ミス・誤照合といった人為ミスが少なくなることである。cは、通信回線を介して遠くの取引先にも帳票が瞬時に送れるため余分な時間がカットできることである。dは、データの転用により、余剰在庫の減少や原価管理など企業経営の効率化が図れる、といったことである。

こうしたEDIによるメリットを実現するために情報交換時の手順の標準化を目指して標準ビジネスプロトコールが策定されることになる。そのビジネスプロトコールにおいて決めるべきこととして、次の6つが考えられている。

- a.情報の定義
- b.構文規則
- c.メッセージフォーマット
- d.標準データコード
- e.伝送方式
- f.メッセージ作成規則

aでは交換を行う情報を定義する。そのためには、業務単位を明らかにし、その個々の業務単位毎にどのような情報単位があるかを定義することになる。どのような文字コードを使うか、メッセージの構造はどうするかを決めるのがbである。aで定義された情報単位に対して、必要な情報項目を考え、標準のフォーマットを作成するのがcである。データ交換においては、種々のコードが用いられる。例えば、企

業・法人などを識別する統一コードが必要であり、それをdでは考える。そして、実際にデータを伝送する際の通信手順・通信プロトコールはeで決める。最後に、fではメッセージ交換の順序や修正・削除の際の手順等が決められる。

こうしたビジネスプロトコールをみると、b, d, e, fはどちらかという物理的な事や通信に関係する事であるが、a, cはデータベース設計、特に論理設計に含まれると考えられる。すなわち、aでの情報の定義は、管理実体型を識別する過程であるし、cでデータ項目を標準化するの管理実体型に対して、その役割を考える過程である。

筆者らはデータベース論理設計の標準データモデルとして構築されているJDMFでの役割概念について、先に考察を行った¹⁾。そこでの所見をビジネスプロトコールのメッセージフォーマット策定に対し試みることにする。

3. 役割の考察¹⁾

本節では、役割の考察を簡単にまとめる。

3. 1 役割の階層性

現在までに考えられてきたいくつかのデータモデルにおいても、役割という概念は導入され、考察されてきた^{5),6),7),8),9),10)}。それらの考察においては役割の共有性、一意性が特性として指摘されていた。JDMFの役割にはさらに階層性という特性を新たに考えることにした。

実体型“注文”を考える。その記述において、“注文の行われる時点”の役割を持った実体型“年月日”を表すため、“注文日”、“注文年月日”などが記述項目として考えられる。それらの項目に共通しているのは、“注文の行われる時点”という役割である。すなわち、“注文時点”とする。

さらに、実体型“出荷”でもその記述で、“出荷の行われる時点”の役割を考えることができるので、先程同様に、“出荷時点”とする。

“注文”、“出荷”は共に人間の行う事柄、すなわち、“行為”である。そして、その“行為”において、“あることが行われる時点”という意味で、例えば、“行為時点”という役割が考えられる。この“行為時点”は“注文時点”、“出荷時点”で表される意味をさらに抽象化されたレベルで表現したものとなる。

“行為時点”という役割は、“出荷”、“注文”に対して、それぞれ“出荷時点”、“注文時点”として現れていると考えることができる。従って、この場合の“行為時点”を基本役割として、その基本役割が“出荷”、“注文”のレベルで“出荷時点”、“注文時点”と派生した形で生じている。これらは、基本役割に対する派生役割とする。そして、それは“出荷”と“注文”を“行為”と抽象化することにより生じている。このように、記述対象となる実体型を抽象化することで、その記述に参加する実体型の役割をそれぞれのレベルで表現できる。これにより、基本役割、派生役割の役割の階層性が生じることになる。

3. 2 概念辞書

記述対象の実体型を抽象化し、そして得られるものを実体素性とする。その実体素性に対して、役割を考えることで役割の階層性が実現される。さらにそれぞれの階層のレベルにおいて成り立つ基本役割・意味的制約条件は、下の階層に対して継承されるとする。

それらをまとめたものを概念辞書とする。特に、意味的制約条件の共有化は、それをチェックするロ

ジックを共有して使えるという点での有効性が考えられる。そして、このような概念辞書はデータベース設計対象を概念的に整理したものとして、情報検索におけるシソーラスと同様な役割を持つと思われる。概念辞書はデータベース設計におけるシソーラスであると考えられる。

4. EDIデータ項目の役割の共有化

前節までに示された階層的な考え方をEDIデータ項目設計に対して適用する。そのために実体素性と役割の階層を原則として最初に与え、それを実際の具体例での実現と比較する。ここでは、〈請求〉を例としてとりあげる。

〈請求〉とは相手が何かを〈注文〉したことより生ずる。そこでまず、〈注文〉について考察してみる。〈注文〉を抽象化し、その上位の実体素性として〈要求〉を考える。

〈注文〉→〈要求〉

〈注文〉の〈要求〉している事柄は〈受渡〉であり、その〈受渡〉は品物の〈納品〉と代金の〈支払〉の2つから成っている。

〈要求〉している事柄…〈受渡〉=〈納品〉+〈支払〉

〈受渡〉→〈交換〉 〈納品〉→〈獲得〉 〈支払〉→〈譲渡〉

〈交換〉=〈譲渡〉+〈獲得〉

ここで、〈受渡〉、〈交換〉のように複数の実体素性から成り立っている実体素性を複合素性とする。なお、ここでの複合素性とは並列または順列の複合のみを考える。複合素性は〈受渡〉のようによく使われるものは辞書には登録しないがまとめて扱うこととする。

〈譲渡〉や〈獲得〉はそれぞれの行為を行う人の立場を変えれば逆に位置づけられる事柄である。従って、それぞれの行為をどちらから見るかという視点を明らかにすることが重要である。この場合は〈注文〉を行う側に視点を置いている。

具体例を考えるときは、“繰り返し”構造を考えねばならない場合が多い。すなわち、〈注文〉は〈単品注文〉ではなく、複数の品物を〈注文〉するし、〈納品〉も一回で〈納品〉せずに〈分割納品〉する場合がある。上記の基本構造は単純なフラットな構造で考えている。“繰り返し”構造はフラットな構造の繰り返しで扱えるし、繰り返し場合は正規化により、実際には別管理実体型として扱われる。従って、ここでは基本としてのフラットな構造で考えることにする。

〈請求〉の基本構造を次のように設定し(図4-1)、その構造を、具体例として財団法人家電製品協会(以下、家電)³⁾と建設産業情報ネットワーク推進協議会(以下、CI-NET)⁴⁾の〈請求〉を表現するメッセージフォーマットを取り上げ比較してみた(図4-2)。なおこの図で示されているデータ項目はすべて各ビジネスプロトコルで必須とされているもののみを取り上げている。

図4-1で注意しなければならないのは、〈商品〉のように上位の2つの実体素性から継承が行われている場合である。このときは、その上位の実体素性で基本役割が同じ役割は一緒になり、上位それぞれの実体素性で異なる役割はそのまま継承される。さらに、〈商品〉の場合、〈集合物〉であることと同時に〈購買物〉であることから、“額”という役割が導出される。これは、〈集合物〉の持つ“数量”役割と〈購買物〉の持つ“価格”役割が同時に存在しうるので、算術操作で導出できるものである。

図4-2で*0は、データの記録に関するデータ項目で<請求>の記述とは直接関係しないものである。*1は、繰り返すものの値を合計したものであるから、これも<請求>の記述とは直接関係しない。*2はいわゆる明細に当たるもので、“繰り返す”項目である。これらは先程述べたようにフラットな構造で考える。

これらのデータ項目に役割を対応付けてみると、図4-3のようになる。なお、*0と*1の項目に対しては直接記述に関係しないので対応付けなかった。

両方において<支払>の“支払者”・“支払相手”に当たるものは、<請求>の“請求相手”・“請求者”と同じであるために、省略されている。また、<請求>においてその“請求場所”が両方において省略されている。このように既に明白である役割は省略される場合がある。

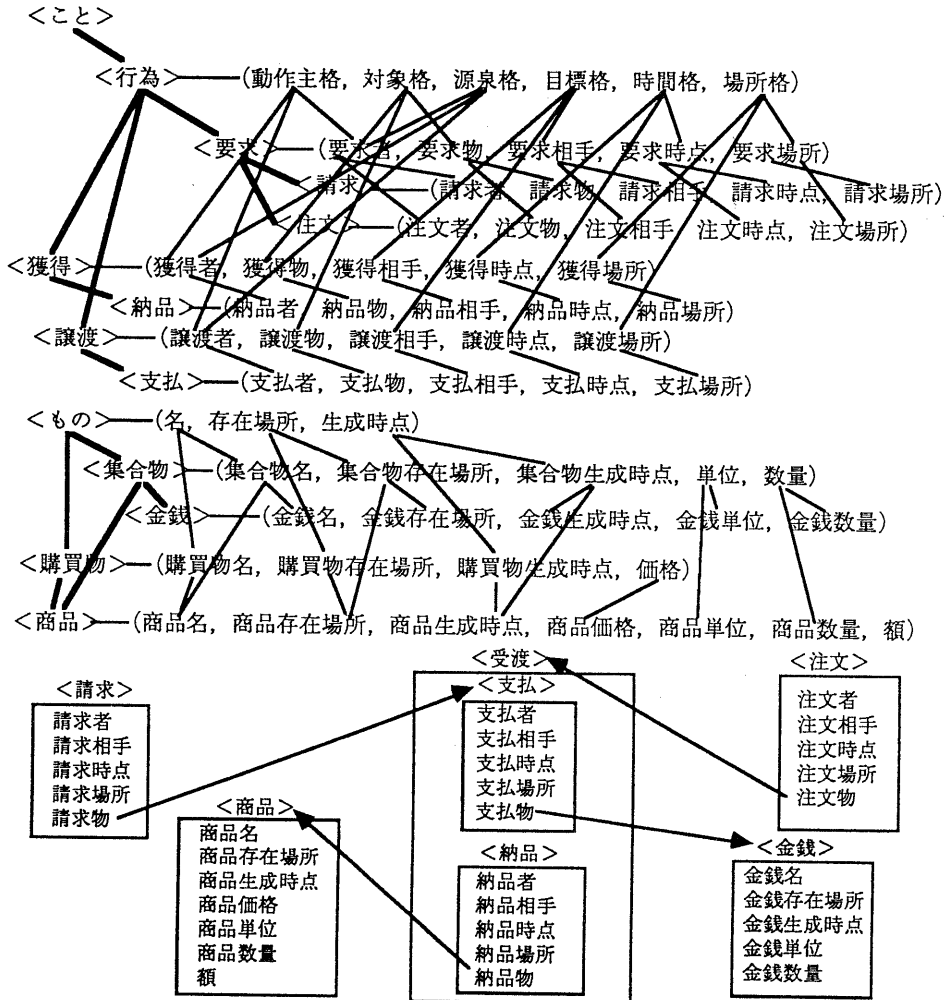


図4-1 <請求>の基本構造

C-I NET・<請求>

*0データ処理No.
*0情報区分コード
*0データ作成年月日
 発注会社コード
 受注会社コード
*0訂正コード
*0帳票No.
 帳票年月日
*0基本参照帳票No.
 基本参照帳票年月日
 受注会社名
 受注会社担当部署名
 受渡先名称
*1注文件名
*1基本数量計
*1基本数量計単位
*1消費税コード
*1基本金額計
*1課税分類コード
*1消費税額
*1帳票金額合計
*0明細コード
*2取引区分コード
*2名称
*2摘要
*2基本数量
*2基本数量単位
*2件別消費税コード
*2単価
*2基本金額

家電・<請求>

*0レコード区分
*0データ区分
*0伝票区分
 発注企業コード
 支払先企業コード
*0データ作成日
 請求月
 伝票日付
*0伝票番号
 納品場所コード
 納品企業コード
*2商品コード識別区分
*2商品コード
*2数量
*2納品単価
*2伝票金額
*1発注企業コード
*1支払先企業コード
*1請求企業コード
*1伝票枚数
*1請求金額

図4-2 <請求>メッセージフォーマット

家電

<請求> : 発注企業コード（“請求相手”）、支払先企業コード（“請求者”）、請求月（“請求時点”）
<納品> : 伝票日付（“納品時点”）、納品場所コード（“納品場所”）、納品企業コード（“納品者”）、商品コード識別区分・商品コード（“商品名”）、納品単価（“商品価格”）、数量（“商品数量”）、伝票金額（“額”）
<支払> : 商品コード識別区分・商品コード（“金銭名”）、伝票金額（“金銭数量”）

C-I NET

<請求> : 発注会社コード（“請求相手”）、受注会社コード（“請求者”）、帳票年月日（“請求時点”）
<納品> : 基本参照帳票年月日（“納品時点”）、受注会社名・受注会社担当部署名（“納品者”）、受渡先名称（“納品相手”）、名称（“商品名”）、基本数量（“商品数量”）、基本数量単位（“商品単位”）、単価（“商品価格”）、基本金額（“額”）
<支払> : 名称（“金銭名”）、基本金額（“金銭数量”）

図4-3 役割の対応付け

図4-3において、CI-NET・家電の2つで異なることとして、

a.CI-NETでは、“取引区分コード”，“摘要”，“件別消費税コード”といった基本構造に対応付けられない項目がある。

b.家電では、＜納品＞の“納品相手”に当たるデータ項目がない。

c.CI-NETでは、＜納品＞の“納品場所”に当たるデータ項目がない。

d.家電では、＜商品＞の記述で“商品単位”に当たるデータ項目がない。

e.CI-NETでは、＜納品＞の“納品者”に当たるデータ項目が2つの項目で表現されている。

f.家電では、＜商品＞、＜金銭＞の“商品名”，“金銭名”に当たるデータ項目が2つの項目で表現されている。

が挙げられる。

e, fは2つの業界でその役割に当たるものをさらに詳しく表現している例になっている。図4-1に示されているのはあくまで基本構造であって、各会社によりさらに細かな表現がされることが当然考えられる。

aもその例となっている。逆にdのように基本構造で設定されている役割が示されない場合もありうる。これには、“単位”が明白であるとの理由から省略されるものと考えられる。

また、bは“請求相手”と“納品相手”が同じものを示すので明白により省略されると考えられる。そして、cも、“納品場所”が明白との理由が考えられる。

結局、この2つの業界において記述の細かさの相違は見られるが、多くの部分においてその役割は共通しており、それは図4-1の基本構造を設定することで理解できる。

概念辞書を予め設定しておくことで、ことばの表現の違いを越えてデータ項目の意味上の統合・共有化が図られる。概念辞書は、業種・業界に寄らないレベル、各業種・業界毎のレベル、各企業レベルの3層での設定が考えられる。先程の例では業種・業界に寄らないレベルでのデータ項目の意味的共有化における有効性が示された。

次に、意味的制約条件における有効性の実例を＜請求＞について見る。ここでは、つぎのような表形式で順に表現していく。

実体素性表	
番号	名
1	こと
2	行為
3	要求
4	譲渡
5	請求
6	注文
7	支払

こと
素性番号
2

行為						
素性番号	動作主格	対象格	源泉格	目標格	時間格	場所格
3	要求者	要求物	要求者	要求相手	要求時点	要求場所
4	譲渡者	譲渡物	譲渡者	譲渡相手	譲渡時点	譲渡場所

“対象格”の指示物が別に存在

要求					
素性番号	要求者	要求物	要求相手	要求時点	要求場所
5	請求者	請求物	請求相手	請求時点	請求場所
6	注文者	注文物	注文相手	注文時点	注文場所

“要求物”の指示物は<支払>素性を持つ

譲渡					
素性番号	譲渡者	譲渡物	譲渡相手	譲渡時点	譲渡場所
7	支払者	支払物	支払相手	支払時点	支払場所

“要求物”の指示物は別に存在し、<支払>素性を持つ

以上の概念辞書構造に対して、2つのEDIのアプリケーションの定義は次のように表現される。

実体型表	
番号	名
#1	請求 (家電)
#2	請求対象支払 (家電)
#3	請求 (CI)
#4	請求対象支払 (CI)
#5	金銭 (家電)
#6	金銭 (CI)

“請求物”は“請求時点”より1月後以降存在していなければならない

<支払>素性を持つ“請求物”が“請求時点より1月後以降別に存在していなければならない

“請求物”は“請求時点”より2月後以降存在していなければならない

<支払>素性を持つ“請求物”が“請求時点より2月後以降別に存在していなければならない

請求					
実体型番号	請求者	請求物	請求相手	請求時点	請求場所
#1	支払先企業コード	#2	発注企業コード	請求月	—
#3	受注会社コード	#4	発注会社コード	帳票年月日	—

支払					
実体型番号	支払者	支払物	支払相手	支払時点	支払場所
#2	発注企業コード	#5	支払先企業コード	—	—
#4	発注会社コード	#6	受注会社コード	—	—

金銭					
実体型番号	金銭名	金銭存在場所	金銭生成時点	金銭単位	金銭数量
#5	—	—	—	—	伝票金額
#6	名称	—	—	—	基本金額

商品コード識別区分・商品コード

上記で示された定義に対して、実際のアプリケーションデータの表はそれぞれ次に示すように表現される。

#1				
主キー	支払先企業コード	#2	発注企業コード	請求月
100	C001	t01	C005	1991.2
200	C002	t02	C008	1991.3
:	:	:	:	:

#2			
主キー	発注企業コード	#5	支払先企業コード
t01	C005	t01_	C001
t02	C008	t02_	C002
:	:	:	:

#5			
主キー	商品コード識別区分	商品コード	伝票金額
t01_	A	G01	50
t02_	B	G03	100
:	:	:	:

#3				
主キー	受注会社コード	#4	発注会社コード	帳票年月日
300	K001	m01	K005	1992.1
400	K002	m02	K008	1992.2
:	:	:	:	:

#4			
主キー	発注会社コード	#6	受注会社コード
m01	K005	m01_	K001
m02	K008	m02_	K002
:	:	:	:

#6		
主キー	名称	基本金額
m01_	鉄材	10000
m02_	サッシ	5000
:	:	:

それぞれのアプリケーションでの<請求>に対して、<支払>の完了のチェックのロジックは、概念辞書の各レベルで定義されている制約条件の継承で構成されていく。ここでは、<要求>レベル・<請求>レベル・各<請求>レベルの3つのレベルで制約条件が記述されている。形式（一階述語）表現を行うと、次のように表される。

xは“家電”における実体型“請求”の実現値であるとする

P(x):xは<請求>素性を持つ

f(x):xの“対象格”の指示物

Q(x):xは<支払>素性を持つ

g(x):xの“時間格”の指示物

$E(x,t):x$ は t 時点以降に存在する	
$\forall x(P(x)\wedge Q(f(x))\rightarrow E(f(x),g(x)+1))$	……最終的に構成される命題
$\forall x(P(x)\rightarrow E(f(x),*))$	……<要求>レベル
$\forall x(P(x)\wedge Q(f(x))\rightarrow E(f(x),*))$	……<請求>レベル
$\forall x(P(x)\wedge Q(f(x))\rightarrow E(f(x),g(x)+1))$	……各<請求>レベル (ここでは“家電”)

上記で $E(f(x),*)$ とはその“対象格”の指示物が存在する時点が<要求>レベル, <請求>レベルでは漠然としている(*)ことを示している。最後の各<請求>レベルでは“家電”とCI-NETで論理式の形は異なってくるが, それまでの形は共通に使える。よって, EDIでのチェックロジックの共有化も可能になりうる。

5. まとめ

実体素性・役割を階層的に表現した概念辞書を考えることで異なる業界で考えられているEDIデータ項目の役割としての意味的な共有化が図られることになる。これにより, 業種・業界に寄らないレベルでの概念辞書の有効性が示された。また意味的制約条件の共有化に対しても有効性がある。

情報システム設計で現れる実体型を抽象化し, 実体素性としてとらえ, その階層構造を与えることでデータの部品化・共有化を概念辞書は目指している。これはオブジェクト指向的なデータ中心のシステム設計と目指すところは同じであり, 具体的に情報システム設計を支援するための概念辞書のプロトタイプ構築を現在構想しつつある。

[謝辞]

本稿をまとめるに当たって, 筑波大学教授穂鷹良介先生から, 多くの有益なご助言をいただきましたことをここに感謝いたします。

[参考文献]

- 1)玉川義人, 穂鷹良介: JDMFにおける役割の考察, データベースシステム研究会, 87, 3, pp.17-26, 1992.
- 2)日本電子機械工業会EDI推進センター: EIAJ取引情報化対応標準1 Cデータ交換規約, 1991.
- 3)財団法人家電製品協会: 家電業界におけるビジネスオートメーションへの対応—ビジネスプロトコールの標準化とネットワーク—, 1990.
- 4)建設産業情報ネットワーク推進協議会: CI-NET標準ビジネスプロトコール (Ver.1.0), 1991.
- 5)E.F.Codd: A relational model of data for large shared data banks, CACM, 13, 6, pp.337-387, 1970.
- 6)P.P.Chen: The entity-relationship model - Toward a unified view of data, TODS, 1, 1, pp.9-36, 1976.
- 7)M.E.Senko, E.B.Altman and P.L.Fehder: Data structures and accessing in data base systems, IBM Syst. J., pp.30-93, 1973.
- 8)H.Biller and B.J.Neuhold: Concepts for the Conceptual Schema, IFIP TC-2, pp.1-30, 1977.
- 9)E.Falkenberg: Concepts for Modelling Information, IFIP TC-2, pp.95-109, 1976.
- 10)C.Bachman and M.Dayna: The role concept in data models, VLDB, pp.464-476, 1977.