

CASE ツール間転送形式の圧縮法

2M-6

森保 健治 伊集院 正
NTT ソフトウェア研究所

1 はじめに

大規模ソフトウェアの開発において、計算機による支援(CASE ツールの利用)は必須になりつつある。しかし、現在市販/流通しているCASE ツールで、全ての開発工程をカバーしているものは少なく、複数のCASE ツールを組合せて使用する傾向にある。

ここで問題となるのは、CASE ツール間の設計情報の共有である。例えば、工程が進む、あるいはあるCASE ツールで扱えない設計を別のCASE ツールで行なうなど、複数のCASE ツールで共通の設計情報を必要とする時、最新のCASE ツールの移行後も過去の設計資産を有効活用したい時などに、以前に利用したCASE ツールの設計情報を再投入するのは禁止的であり、CASE ツール間で設計情報を転送することが有効である。

EIA(Electronic Industries Association)では、CASE ツール間情報交換用標準形式としてCDIF[1]を提案し、現在国際標準化のための検討が進められている。CDIFの特徴としては以下が挙げられる。

- (特徴1) 幅広い設計情報カテゴリ (Subject Area) をサポート
- (特徴2) 転送内容と転送形式の分離
- (特徴3) 意味情報と表示情報が転送可能
- (特徴4) 人が読んで理解できる転送形式

上記の特徴を持つCDIFであるが、実際に使う場合、(特徴3)(特徴4)のために非常に転送データ量が大きくなってしまふことが問題となる。本稿では、(特徴2)を生かして、簡単な変換により転送データ量を50%以上削減できる可能性を確認したので、報告する。

2 CDIF 転送形式

CDIFでは、転送内容としてDF図、ER図、DS図など多くの種類の設計情報を扱うことができ(特徴1)、それを転送形式に変換して、CASE ツール間で交換する。図2に、図1のER図を標準で用意されたエンコード方式[5]で変換した転送形式の一部を示す。

3 エンコード方式と転送データ量

図2に示すように、オブジェクトおよびオブジェクト間の関連の宣言は、そのまま読み下すことができる形式である(特徴4)。しかし、図1のような簡単な図におい

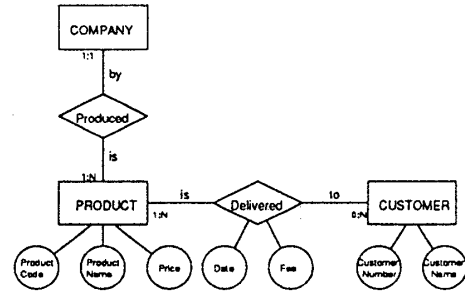


図1: ER図の例

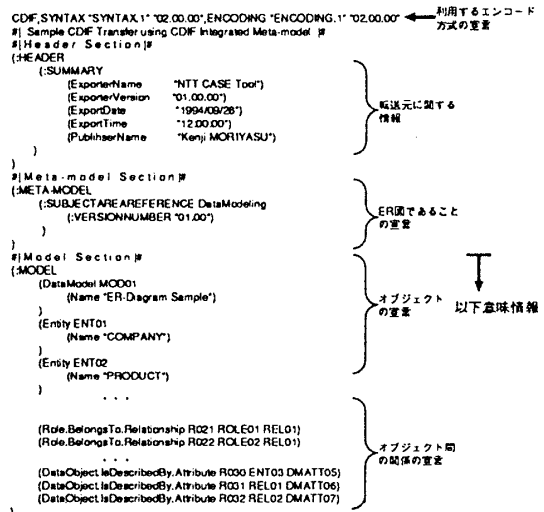


図2: ER図の転送形式の例

ても、転送データ量は大変大きなものとなっている(意味情報のみの転送で2866byte、もし表示情報も含めると、この10倍になると予想される)。

転送データ量の削減方法を考える上で、ここではCDIFの使い方を考察する。

一般に、あるソフトウェア開発に使用されるCASE ツール群が共通的に必要とする設計情報は、一つのプロジェクトに閉じていることが多い。つまり、CDIFは、一つのプロジェクトで使用するCASE ツール間で設計情報を交換するのに利用されることになる。

一方、(特徴2)により、交換される設計情報のエンコード方式は、転送内容とは独立している。したがって、転送元CASE ツールと転送先CASE ツール間の同意が取れば、エンコード方式を工夫することにより、転送データ量は削減できることが期待される(ただし、現在は、1つの標準が規定されている[5])。前述のような同一プロジェクト内でのデータ交換では、この同意が取れることが多いと考えられるため、転送データ量を削

¹A Reduction Method of Amount of Transfer Information between CASE Tools
Kenji MORIYASU, and Sei IJUIN
NTT Software Laboratories

減するエンコード方式を定めることが可能である。

4 転送データ量の圧縮

4.1 転送データ量の圧縮法

転送データ量の多さは、各 Subject Area で規定された、人間がテキストとして可読なキーワードをそのまま転送していることが大きな要因である(特徴4)。しかし、転送されるデータは、CASE ツールが理解できればよく、そのままの形で人間が可読である必要はあまりない。したがって、このキーワードを短縮されたキーワードに置き換えるようなエンコード方法を取ることが簡単かつ有効であると考えられる。

このキーワードの置き換えには、以下のルールが守られることが必要である。

- 置き換えるキーワードは1つの Subject Area 内で重ならないようにする
- Subject Area は、共用して使うことができるので、同時に使用する Subject Area 間(例えば、ER 図と DS 図)でもキーワードは重ならない必要がある。

表1に置き換え例を示す。

表1: 置き換え例

キーワード	短縮キーワード
Attribute	A
Entity	E
DFMProcessDefinition	DFMP
FlowDefinition	FD
...	...

4.2 圧縮の効果

本稿で行なった実験では、転送データのうち意味情報のみを対象とした。最も利用されている図(DF 図, ER 図, DS 図)の簡単な例(図3, 図1, 図4)について、それぞれオリジナルのエンコード法と短縮するエンコード法によって変換された転送形式のデータ量を比較したものを、表2に示す。

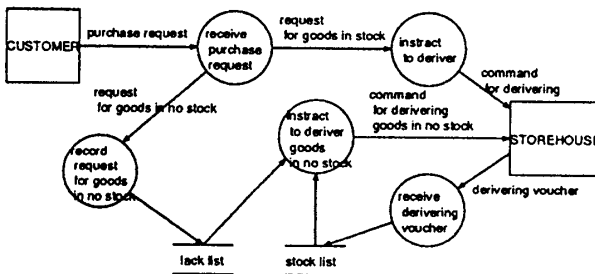


図3: DF 図の例

表2に示すように、ほぼ50%から60%程度に圧縮できることが分る。図の種類によって圧縮効果が異なるのは、Subject Area 毎にキーワード長の平均が違うためである。また、この圧縮効果は、各行でほぼ一定な

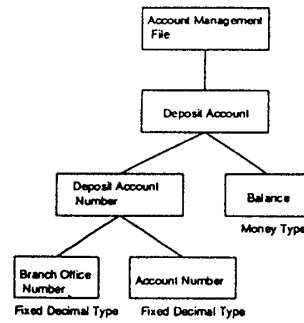


図4: DS 図の例

表2: 転送データ量の比較

	DF 図	ER 図	DS 図
オリジナル: A	5739	2397	2092
本方式: B	3019	1496	1152
圧縮効果: B/A (%)	52.6	62.4	55.1

め、図の複雑さ(複雑なほど行数が増える)にはあまり影響がないと思われる。

キーワードの選び方(例えば、出現頻度が高いキーワードの長さは、より短くする、など)、および各オブジェクトの識別子(図2の DMATT07 など)の選び方によっては、圧縮結果を50%以下とすることも可能であると思われる。

5 おわりに

本稿は、CDIF が転送内容と転送形式が分離していることに着目し、送り側と受け側の同意が取れる環境において、エンコード方式を工夫すれば、転送データ量の大幅な圧縮が可能であることを示した。この圧縮法は、以下の特徴を持つ。

- 転送内容には影響を与えない。
- 簡単なキーワードの置換だけなので、エンコード/デコードに掛る計算量は少ない。
- 将来キーワードが増えても、置換のための表が増えるだけである。

参考文献

- [1] CDIF - Overview(Ver.7)
- [2] CDIF - Integrated Meta-model Data Flow Model Subject Area(Ver.7)1994.3
- [3] CDIF - Integrate Meta-model Data Modeling Subject Area(Ver.5) 1993.10
- [4] CDIF - Integrated Meta-model Data Definition Subject Area(Ver.8) 1994.3
- [5] CDIF - Transfer Format - Encoding ENCODING.1(Ver.7) 1993.10