

データ分析支援ツール“DBSOLVER”（3）

7G-4

—ER図統合支援機能—

山口陽一 金枝上敦史† 鈴木由美子† 菅原幹夫 小林博
 三菱電機（株） 情報システム製作所 †情報システム研究所

1. はじめに

近年のダウンサイジングやRDBの普及の流れの中で、システム開発におけるデータ中心アプローチ（DOA）の重要性が高まってきている。ここに報告するツールは、このDOAを実現するためのデータ分析作業を支援するものである。

本ツールは、分析対象の業務で扱われるデータ項目を抽出し、そこからシステム全体のデータモデルを組み上げるボトムアップのデータ分析を支援する。本ツールは、大きく分けて**標準化、正規化、ER図統合**の3つの支援機能からなる[1]。本稿では、本ツールの特長的機能であるER図統合支援機能について報告する。

2. 統合とは

ボトムアップのデータ分析では、対象となる業務で扱われる伝票、帳票、現行システムの画面又はファイル定義等（これらをフォームと呼ぶ）を基に、正規化手法を用いて各フォーム毎のER図（フォームER図と呼ぶ）を得る[2]。こうして得たフォームER図をシステム全体のデータモデルを表す一つのER図（全体ER図と呼ぶ）にまとめあげる作業を『統合』という。

図1に本ツールにおけるER図の統合の様子を示す。本ツールでは、フォームER図を全体ER図に順次流し込むようにして統合を行う。既にいくつかのフォームER図を統合した全体ER図と他のフォームER図には、互いに一致するエンティティ及びリレーションシップが現

れる。これらの一致の発見とその一本化が統合の主たる作業で、本ツールはこれを支援するために、以下のような特長的機能を備える。

- ・類似エンティティ検索機能
- ・標準データ項目による属性併合機能
- ・エンティティ統合順序誘導機能
- ・類似/冗長リレーションシップ検出機能

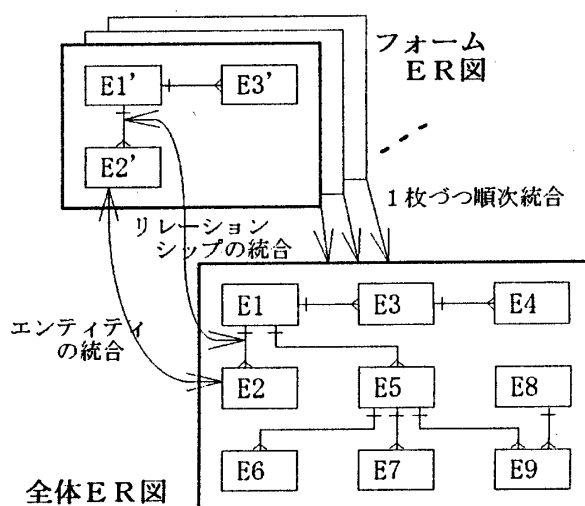


図1. ER図の統合

3. 類似エンティティ検索機能

類似エンティティ検索機能は、全体ER図からフォームER図上のエンティティと一致するものを検索するものである。本ツールでは、エンティティの主キー属性の比較検索、エンティティ名のキーワード検索等が可能である。主キー属性の比較は、属性のタイプ、桁数、ドメイン、標準データ項目、及び名称の部分一致といった比較条件をそれぞれ独立に設定できる。これらの検索条件を適宜組合せて、全体ER図から一致するエンティティの候補を絞り込むことができる。

4. エンティティの統合

エンティティの統合は、エンティティ名等の決

Tools for Data Analysis : DBSOLVER (3) -computer aided synthesis of entities and relationships-
 Youichi YAMAGUCHI, Atsushi KANAEGAMI, Yumiko SUZUKI, Mikio SUGAWARA, Hiroshi KOBAYASHI
 MITSUBISHI ELECTRIC CORPORATION

定と属性（データ項目）の併合によって行われる。属性の併合に際しては、統合する2つのエンティティの全ての属性の一致不一致を分析者の判断で決定する。本ツールでは、標準化機能と連携し、同一の標準データ項目を割り当てることで属性の一致を表わすようにした。この標準データ項目による属性併合機能により、事前に標準化が完了している属性に対する併合作業を省略できる。

この属性の併合に際して、リレーションシップにより誘導される属性（外部キー属性）に対する考慮が必要である。図2は、ER図と外部キー属性の例で、エンティティAとそれに従属するエンティティB、その従属を表わすリレーションシップ、及び各エンティティの主キーを構成する属性のセットを示してある。エンティティBには、その主キー属性にエンティティAの主キー属性（属性a）が含まれる。本来このような外部キー属性は、上位にあるエンティティA側の統合時に一致不一致を判定されるべきである。このため本ツールでは、エンティティ統合順序誘導機能を設け、上位のエンティティから順に統合を行うよう誘導する。

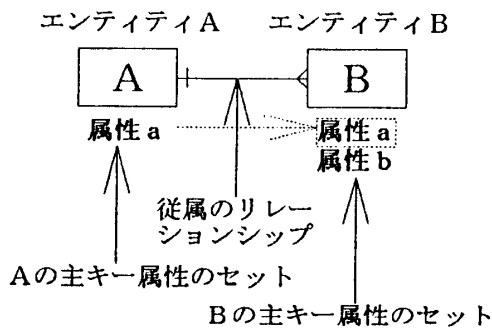


図2. 従属するエンティティの主キー属性

5. リレーションシップの統合

フォームER図から全体ER図にエンティティが一つずつ取り込まれるつど、その間のリレーションシップも全体ER図に取り込む。この際、同一のリレーションシップがあれば適宜一つに統合しなくてはならない。中でも、親子関係（従属関係）を表わすリレーションシップは、子側エン

ティティの主キーを構成し、その子側エンティティの意味を左右するため、特に注意を要する。

図3は、リレーションシップの統合の例である。A-D間のリレーションシップに着目すると、エンティティの意味を保持するためには、フォームER図と全体ER図にある二つを、その意味を検討した上で統合する必要があることがわかる。また同様に、フォームER図におけるA-C間のリレーションシップは（正規化時にエンティティBを見落としたものであるが）、統合後ER図には存在すべきでない。本ツールは、この例のようなケースを誘導される外部キー属性を基に検出する類似/冗長リレーションシップ検出機能を有し、分析者に処置を促す。

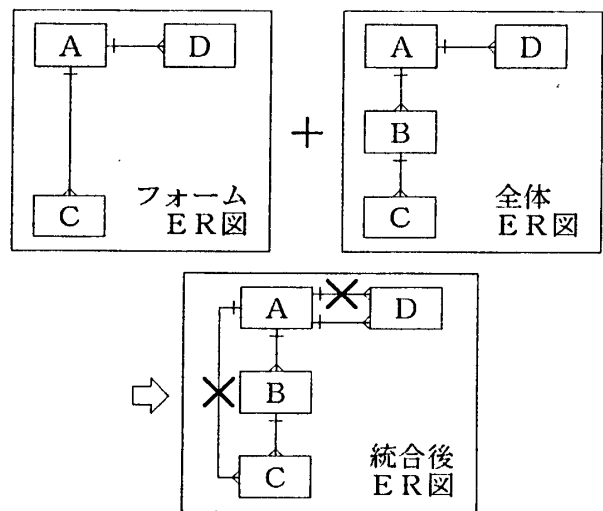


図3. リレーションシップの統合

6. おわりに

本稿では、データ分析支援ツール DBSOLVERの特長的な機能であるER図統合支援機能について報告した。現段階ではまだルーチンワークを支援するに留まるが、今後の実分析作業への適用の中で、更なる機能拡張を図りたいと考えている。

参考文献

- [1] 金枝上, 他, データ分析支援ツール“DBSOLVER”(1) —狙いと機能概要—, 情報処理学会第50回全国大会, 7G-2
- [2] 鈴木, 他, データ分析支援ツール“DBSOLVER”(2) —分析の流れ—, 情報処理学会第50回全国大会, 7G-3