

統合的データモデリング支援環境“Darwin”(3)

4 E - 6

—論理データモデル支援ツール“Darwin/LD”の紹介—

武田 和彦 郡司 賢 丹羽 隆 加藤 雅樹



NTTデータ通信株式会社 技術開発本部

1 はじめに

近年のデータ中心アプローチ(DOA)によるシステム開発の普及の中で、データモデリングにおける概念的・論理的分析のフェーズは、“単なる”データベース設計の前段階といった“軽い”意味合いから、先進的な戦略情報システム構築のために企業のもつ全情報を掘起こし、効率よくそれらを利用するための青写真を作成するといった最も重要なステップであると認識されつつある。

本稿では、統合的データモデリング支援環境“Darwin”のうち、論理データモデリングのフェーズを支援するためのツールである“Darwin/LD”のコンセプト及び機能概略について説明する。

2 コンセプト

論理データモデリングの目的は、分析対象領域の事象を抽象的に捉えてモデル化し、データの物理的な格納方法やアクセス手段とは独立にデータの概念的な構成を記述した論理データスキーマを導出することにある。また標準化、正規化等の作業を繰返し行うことにより、冗長性のない論理的に矛盾のないスキーマを導くことにある。

通常、論理データモデリングは、複数のデータ分析者(Data Analyst)と呼ばれる人々の協調作業によってなされる。“Darwin/LD”のコンセプトを決定するに当たり、実際に、当社において、DOAによるシステム開発を実施している部門のデータ分析者の人々にインタビューを行った。この結果を受け、以下に述べるワークフローを仮定した。

論理データモデリングでは、まず始めに、1つのデータモデルを細分化した単位であるサブジェクトエリアを定義する。このサブジェクトエリアが1分析者当たりの基本的なワークスペースとなる。ここに分析対象領域のERダイアグラムを記述していく。このERダイアグラム上に定義されたエンティティ・アトリビュートといったオブジェクトに対し、用語辞書を用いた名称標準化を行い異音同義語等のデータを排除すると共に、リレーションシップにおいては推移的關係にあるものなど冗長なリレーションシップを排除する。また正規化規則に従って第1～第3正規化を実施し、データモデルの洗練をはかる。最後にデータモデルに対し、数件のチェック項目を定義し、モデルの品質を検査する。

以上のワークフロー分析から、“Darwin/LD”では、こうした一連のデータ分析者の作業を支援し、論理データモデリングが効率よく実施できることを第1のコンセプトとした。

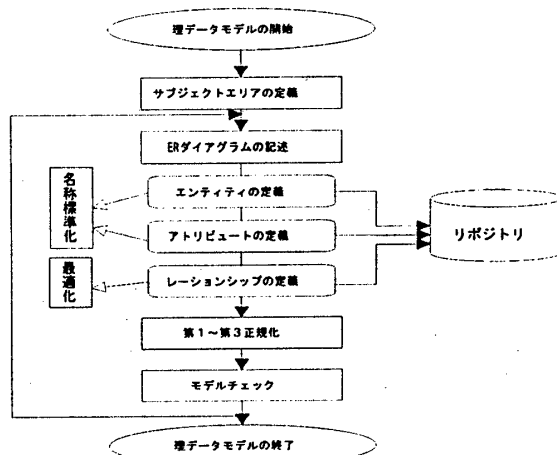


図1 論理データモデリングのワークフロー

3 機能概要

3.1 ERダイアグラムエディタによる視覚的編集

“Darwin/LD”の有するERダイアグラムエディタでは、以下に示すERモデルを表現することが可能である。

- 単純関係、 ●再帰関係、
- 複雑関係、 ●エンティティ階層

ERダイアグラムエディタは本ツールにおける標準的な作業空間であり、高度なドローイング機能、GUIによる直感的でわかり易いユーザインターフェイスを提供している。

また、次に挙げる“一覧表形式入力”と併用することで2つの異なった視点からの編集作業を行うことができ、効率的なモデリング作業が可能である。

3.2 一覧表形式による簡易な設計情報編集

一覧表形式ウィンドウでは、設計詳細情報の一括入力が可能である。従来のCASEツールにありがちであったダイアログによる設計情報個々の単体入力に比べ、はるかに見通しの良い効率的な作業環境を提供している。

一覧表形式ウィンドウと、ERダイアグラムエディタは、互いにメッセージパッシングを行うことで表示情報の同時更新を行う。

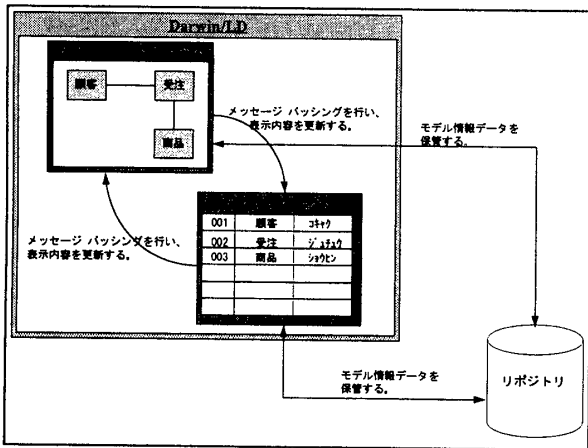


図2 "Darwin/LD"におけるウィンドウ間の連動

3.3 サブジェクトエリアによる分散設計支援

対象となるシステム全体を表す1つの論理データモデルを、サーバー上のリポジトリに格納する。クライアント側では、この1つの論理データモデルを複数のサブジェクトエリアに細分化し、モデルの一部に対する編集を行う。この仕組みにより、複数分析者による並行作業を行うことが可能である。

3.4 名称標準化によるデータ重複の徹底的な排除

サーバ上のリポジトリには用語辞書が実装されており、アトリビュート、エンティティ、ドメインといったオブジェクトに対して名称標準化を行い異音同義語等の冗長なデータの排除を行う。名称標準化の規則としてDurellの提唱した名称標準化規則を拡張^{[1][2]}したものを採用した。下図に名称標準化の規則を記述する。

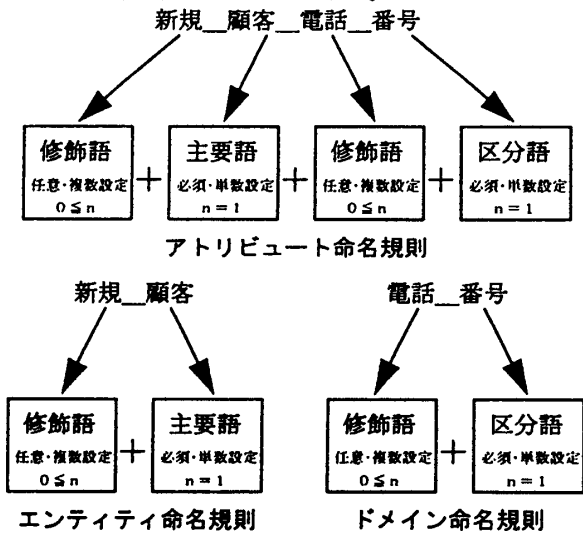


図3 名称標準化規則

3.5 ドメインによるアトリビュート値の一括管理

ドメインとは、アトリビュートの値がとれる範囲を規定したものである。1つのアトリビュートは必ず1つのドメインと結び付けられなければならない。

ドメインでは、具体的にはデータ型、データ長、許容値等、そのドメインと結び付けられているアトリビュートの値の範囲を限定する制約事項を定義する。また、ドメインは複数のアトリビュートで共有されるものであり、

1つのモデルにおいて、共通に利用されるオブジェクトであるので、ドメイン辞書としてリポジトリ上に実装した。

3.6 第1～第3正規化作業を支援

正規化を行うためのダイアログを用意し、正規化作業を簡易に行えるよう支援している。正規化によって新規に生成されたエンティティは自動的にダイアグラム上に配置され、また一覧表形式ウィンドウへも挿入される。

3.7 リレーションシップの最適設計のための支援

冗長リレーションシップの検索・排除や、多対多リレーションシップの検索・排除等の機能を有する。

冗長リレーションシップの検索では、特に下図に示すような推移的な関係にあるリレーションシップを発見する。また、RDBMSへのフォワードを考慮して外部キーの自動生成及び参照整合性オプションの設定も可能となっている。

3.8 論理データモデルの妥当性チェック機能

20数個のチェック項目を設定し、設計されたモデルが論理データモデルとして妥当かどうかを診断する。

3.9 インポート・エクスポート機能

他のCASEツールとの連動を考慮し、テキストファイルによる設計情報の搬入・搬出機能を有する。尚、テキストファイルのフォーマットは将来的にはCDIFの規格を採用する予定である。

3.10 フォワード/リバースエンジニアリング機能

"Darwin/LD"によって設計された論理データモデルは、リポジトリに格納され、次工程のツールである物理データベース設計ツール"Darwin/PD"にフォワードエンジニアリングされる。逆に、"Darwin/PD"によって設計された物理データモデルは、"Darwin/LD"にリバースエンジニアリングでき、既存のデータベース資源の再利用が可能である。

4 おわりに

本稿では、論理データモデル支援ツール"Darwin/LD"のコンセプトと機能概要について報告した。

今後は、本ツールを実プロジェクトへ適用し、ブラッシュアップをはかるとともに、以下に示す機能拡張を予定している。

- 分散開発支援環境の更なる充実をはかり、より高度なグループウェア環境への対応。
- GUIインターフェイス作成ツールとの連動。
- ERダイアグラムに描画最適化能力を付加。
- リポジトリにOODBMSを使用。
- OODBMS向けのモデリング支援環境への発展。

参考文献

[1] 加藤 雅樹：Durell命名規則の拡張によるデータ項目名称付与の自動化，情報処理学会第49回全国大会，3V-07，1994
 [2] Durell.R.W：データ資源管理，日経マグロウヒル，1987