

DB設計支援ツールDBpromptのアーキテクチャ

7H-6

大久保成隆 町原宏毅 関根純 中川優
NTT情報通信網研究所

1. はじめに

DBシステムの開発には、DB化するデータの分析段階から設計、構築までの一連の作業工程で多岐にわたる技術、知識が要求される。

また、近年の計算機システムの増加、ユーザニーズの多様化を背景として複数システム間でデータの流通を行うための技術、既存システムの更改に対応する技術が重要となってきている。

そこで、筆者らは既存DBの再利用、データの分析、標準化、DBの設計、構築の各作業工程を体系的に支援するツールDBpromptを開発した。

その結果、DB設計支援ツールでは以下の技術が重要であることを確認した。

- ①データ標準化ツールを含むツール間の連動技術
- ②システム更改におけるDB設計のリバースエンジニアリング技術
- ③DBシステムの性能設計技術

本報告ではDB設計支援ツールで実現している上記技術の具体的内容について述べる。

2章でDBpromptの構成とその基本機能について、3章、4章ではDBpromptの重要な機能であるリバースエンジニアリング機能、性能設計機能のそれぞれについて述べる。第5章ではDBpromptを用いてDB設計を行う場合の利用手順を示す。

2. DBpromptの構成と基本機能

DBpromptはDB設計の各作業行程を支援するツール群から構成される(図1)。

DBpromptを構成している各ツールは情報資源辞書管理システムIRDSの機能を用いて、共通の設計情報辞書へアクセスできる。

各ツールが実現している機能を表1に示す。

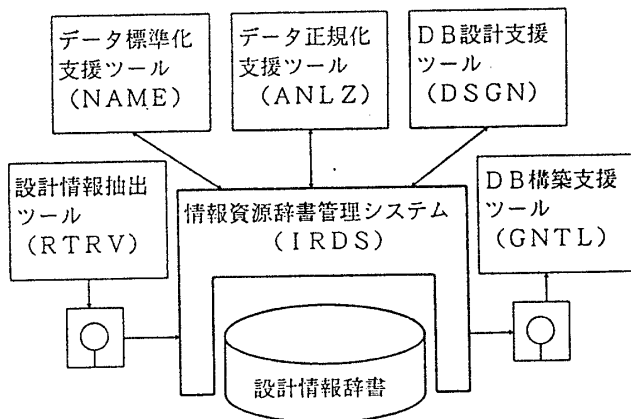


図1 DBpromptの構成

Architecture of database design tool DBprompt
Naritaka OKUBO, Hiroki MACHIHARA, Jun SEKINE,
Masaru NAKAGAWA
Network Information Systems Labs.,NTT

表1 DB設計支援ツールの機能

ツール名称	機能概要
設計情報抽出ツール (RTRV)	①ワープロテキストとして作成された設計帳票からの設計情報抽出 ②コード変換
データ標準化支援ツール (NAME)	名称、コード体系の統一を支援する。 ①規則に合わない名称のチェック ②各システムの名称、コードの一元管理 ③類似データの分類
データ正規化支援ツール (ANLZ)	DB概念設計を支援する。 ①データ間の関係矛盾のチェック ②正規化
DB設計支援ツール (DSGN)	DB論理設計を支援する。 ①DB設計情報のチェック ②DB容量の見積もり ③性能診断 ④設計帳票、ワークシートのレポート
DB構築支援ツール (GNTL)	DB物理設計を支援する。 ①ファイル容量の見積もり ②メモリ容量見積もり ③DB構築用定義コマンドの作成

DB設計ツールの連動については既にいくつかの製品が実現しているがデータ標準化ツールとの連動を可能としているものはない。

DBpromptにおいては、データ標準化ツールと連動させることにより、DB設計の情報をそのままデータ標準化に使用できるため、標準化の漏れの削減及び、稼働の削減が可能になった。

3. リバースエンジニアリング機能

DBpromptでは、既に設計済みのDB設計結果を新規DB設計あるいはDB再設計に利用するためのリバースエンジニアリング機能を実現している。

DB設計におけるリバースエンジニアリング機能については、既にDBスキーマからDB定義情報の抽出を行うツールが実用化されているが、DB設計情報として重要な以下の項目がDBスキーマには存在しないという問題がある。

- 1) データ項目の定義名以外の名称(日本語名等)
- 2) データ項目の意味、内容

筆者らは上記の設計情報の多くが、システム設計時にワープロ文書として作成される設計帳票に記述されることに着目し、設計帳票からの設計情報抽出機能を実現した。

リバースエンジニアリング機能の動作概要を図2に示す。リバースエンジニアリング機能は制御情報の指定内容に従って、設計帳票から情報を取り出し、設計情報辞書への格納を行う。表2に示すように設計帳票の物理フォーマット、論理フォーマットにより、抽出箇所の指定を行い、特定の帳票フォーマットや特定のシステムに依存しない汎用的な設計情報抽出を可能としている。

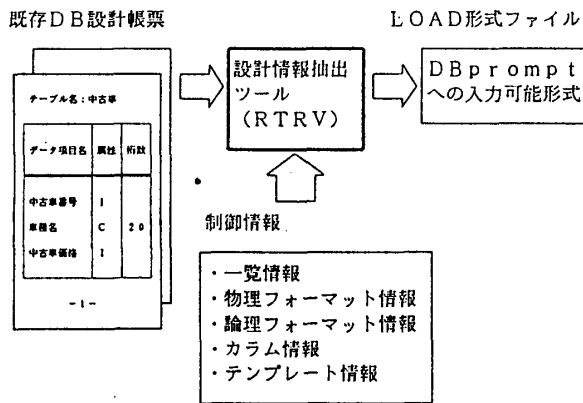


図2 リバースエンジニアリング機能の処理概要

表2 リバースエンジニアリング機能の制御情報

制御情報	内容
一覧情報	設計帳票ファイルのファイル名を指定する。
物理フォーマット情報	設計帳票のフォーマット形式と抽出箇所を指定する。
論理フォーマット情報	物理フォーマット情報で指定した設計情報の階層関係(上位項目に対する繰り返し)を指定する。
カラム情報	設計情報を格納する設計情報辞書のカラムの属性を記述する。
テンプレート情報	設計帳票から抽出する設計情報と設計情報辞書との対応関係やデフォルト値の指定を行う。

4. DB性能設計機能

DB設計支援ツールとしてはデータの正規化ツールのようなデータの特性のみを考慮したものが実用化されているが、実際のDB設計では例えば、

- ・テーブルアクセス時の長時間待ちを削減するため、負荷が大きいテーブルは分解する
- ・テーブルアクセスのI/O回数削減のため、同時にアクセスする可能性の高いテーブルは統合するという業務処理を考慮した設計が行われている。

そのため、DB設計支援ツールには業務の性能条件を考慮した機能が必要となる。筆者らは設計情報辞書に業務のDBアクセス情報を持たせ、業務のアクセス情報からDB処理まわりの性能診断を行うことでDB性能設計機能を実現した。更に、性能診断に関する知識を適宜付加可能なシステム構成を実現した。

・性能設計機能の例を以下に示す。

1) テーブル統合候補の抽出

任意のテーブルについて、そのテーブルをアクセスしている全業務のトラフィックのうち特定のテーブルをアクセスしているトラフィックの割合を調べ、割合が高ければそれらのテーブルをテーブル統合候補としてレポートする。

2) テーブル分解候補の抽出

テーブル毎のトラフィックの負荷を評価し、負荷の高いテーブルを分解候補としてレポートする。

3) インデックス設計の診断

① インデックス削除候補の抽出

インデックスに指定されている項目の更新は性能劣化要因となるため、そのような更新を行っている業務の存在を調べ、更新頻度の高いインデックスを削除候補としてレポートする。

② 業務の検索条件による診断

インデックスを有効に使用していない業務の存在を調べ、その検索条件からインデックスの変更または付与候補をレポートする。レポート例を図3に示す。

<業務の検索条件によるインデックス診断>

--- 6. 社員情報テーブル ---
社員名簿作成業務がインデックスページ総名処理により絞り込んだ範囲でデータベースの評価を行っています。
以下に示す列で構成されるインデックスの付与を検討してください。

列番号, 列名

2. 社員名
3. 性別
5. 年齢

図3 性能診断の出力例

5. DBpromptの利用手順

今回、報告したDBpromptの利用手順をシステム更改時の例で示す(図4)。

ここでは以下を行う。

- 1) リバースエンジニアリング機能による設計情報の抽出
 - 2) 性能設計機能による新規DB設計
 - 3) データ標準化機能による規約違反のデータ項目名チェックと標準名生成
 - 4) DB構築支援ツールによるDB構築
- 実システム開発では、データの標準化とDBの設計、構築担当が異なる場合が多く、図4のようにデータ管理者とDB設計者の作業は分担して行われる。また、この手順のようにデータ標準化を設計終了後に行うことで設計途中で発生するDB情報の追加、変更による標準化作業の手戻りをなくすることができる。

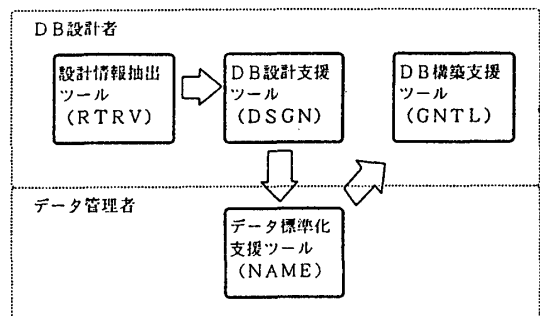


図4 DBpromptの利用手順

6. まとめ

本稿では、DB設計支援ツールDBpromptを構成している主要な技術であるデータ標準化、リバースエンジニアリング、性能設計について述べ、その有効性について確認した。

DBpromptはツールの連動により、これらの機能が必要最小限の入力で実現できるため、実用性に優れたツールであるといえる。