

化学反応設計支援システム用知識ベースの設計

5M-4

藤原 譲¹, 大保 信夫¹, 中山 堯², 鈴木 功¹

1. 筑波大学, 2. 国際科学振興財団

1. はじめに

化学反応設計システムにおいては、多くの化学反応の事例を集めたデータベースと、反応規則、化合物の性質、反応の一般的性質を記述した知識ベースを有機的に結合することが重要である。本報告では、大量の事例データベースと知識ベースを統一的に表現、管理することを目的としたEDTM (Evolutional and Dynamic Table Model) を基にした化学反応設計支援システムに関する基本設計概念を述べる。

2. 化学反応設計支援システムの概要

化学反応設計システムにおいては以下の二点が重要である。

1) 極めて数多くの化学反応の事例が既に報告されており、その多くが複雑な反応条件等のため、反応規則の単純な適用からは推論不可能であるか、推論可能としても莫大な計算

量を必要とする。これらの反応事例を設計に取り込むことなしに有効な支援システムを構築することは困難である。

2) LSI設計等の場合と異なり、設計対象に関する全ての知識を表現しておくことはまず不可能である。従って、データ、知識とも不完全な情報が与えられ、必要に応じてそれらの情報から新しいデータ、知識を抽出して利用することが不可欠となる。言い換えると、反応設計システムは本質的に動的な知識体系を有していなければならない。即ち、設計者の帰納、類推、直観等を十分に支援し、それに基づき自らの知識体系を動的に更新可能なものでなければならず、従来の知識工学の中心である演繹的推論は以上の目的のための局所的なルーチン・ワークと考えられる。以上の点を考慮した化学反応設計支援システムの概要を図1に示した。

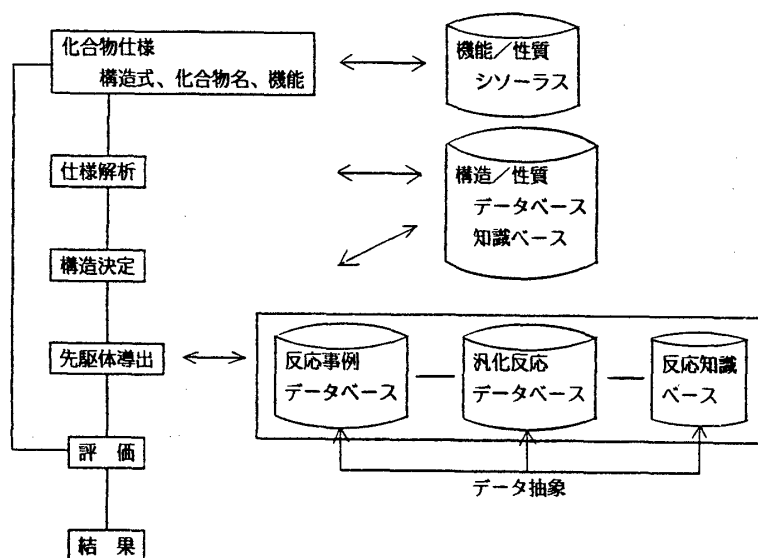


図1. 化学反応設計支援システムの概要

Design of Knowledge Base for Supporting Design of Chemical Reactions
 Yuzuru FUJIWARA¹, Nobuo OHBO¹, Takashi NAKAYAMA², Isao SUZUKI¹
 1. University of Tsukuba 2. International Foundation for Advancement of Science

本システムは、目標物質の仕様が入力されると、それを解析し目標物質の構造決定を行い、それに至る反応系列の推論を行う。その中心となるのはプロダクション型規則で表わされる反応規則の集合、試薬、触媒等に関する知識ベース、Organic Chemistry全63巻から成る反応事例データベースと、それを抽象化した汎化反応データベースである。設計者の帰納、類推、直観等に基づき個別の反応事例の集まりから、それらをより抽象的な反応へ体系化してゆくためのツール群とデータベースの再構成機能がシステムの中核となる。

3. EDTM

EDTMは、実行効率、管理安定性（特に共有上の諸問題点）、論理問合せに対する適合性、既存のデータベース技術の利用可能性等を考慮して関係データベースをその基本とし、それに以下の機能を付加している。

1) 複雑なデータ型支援：化学グラフ、反応規則、時間等複雑な意味の附随するデータ型を支援するため、EDTMでは全てをオブジェクトとみなし、表、行、列の手続き的知識をデータ管理部門に付加することを可能としている。実際には、関係データベースのエントリは全てオブジェクト識別子であり、オブジェクトの各定義は別にオブジェクト定義ファイル群に分離格納・管理される。データベース操作は、オブジェクト識別子を通常の文字列とみなした場合の関係演算操作と、オブジェクト定義部のメソッドに依存する演算群に分かれる。

2) 意味論的モデリングの支援：汎化、統合等を基本とする意味論的モデリングおよびそれに関連するデフォルト、遺伝等のヒューリスティックな推論を実行するため、フレーム型知識表現との結合を可能とする関係データベース中にフレーム・インスタンスに対するオブジェクトをフレーム中に関係問合せに対するオブジェクトを許すことにより、オブジェクト定義部のメソッドを介して関係データベースとフレームを結合している。（この概念図を図2に示した）

3) メタ・データ・アクセス機構：反応設計支援システムでは、データベースの再構成が極めて頻繁に生ずる。そのためには、ユーザからデータに関するデータ（メタ・データ）を効率的にアクセス、処理する機能が不可欠である。EDTMでは関係名、属性名等は構文的には通常的应用データと殆ど同様に扱うことが可能である。これらの演算に関する意味は、オブジェクト定義部の関係型オブジェクト中のメソッドで定義される。

4. むすび

本システムはデータ入力、および表現においても多くの問題を抱えており、システムの現状とともにデータの現状と課題についても述べる。

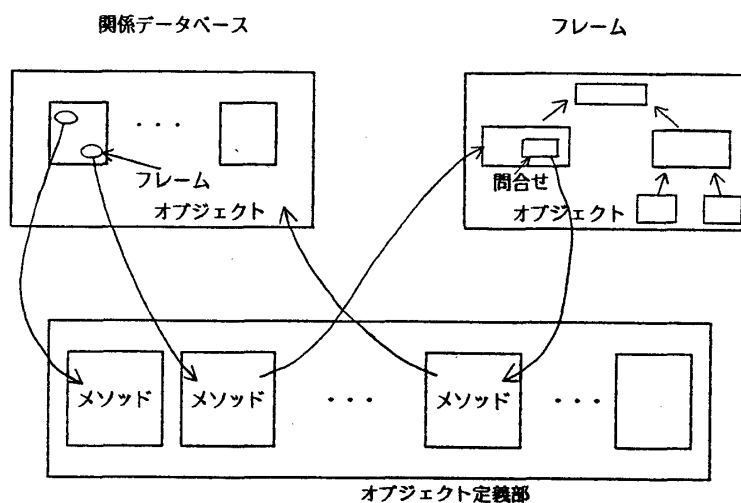


図2. 関係データベースとフレーム