

有機合成反応データベース管理システムの開発とその高次利用に関する研究

材料システム工学専攻 845501

遠藤 智明

本研究では、化学の領域における計算機を利用するための新しい試みである、化学的知識を計算機により自動的に獲得するための一連の研究を行った。ここでは、知識を取り扱う対象を有機合成反応におき、有機合成反応に関する知識を自動獲得する手法を開発した。具体的には、1)反応データベース管理システムの開発、2)反応中心の自動認識機能の開発、3)反応経路列挙機能の開発、4)反応における単位知識の自動獲得機能とその利用プログラムの4つの核となる機能を開発した。反応データベース管理システムにおいては、個別反応における反応情報の登録・検索を行ない、反応に関する出発物、生成物の構造情報とともに、反応条件、反応の収率などの情報と文献情報を登録できる機能に加えて、個々の化合物に対する沸点、融点などの物性情報の管理を行える。反応検索の機能としては、反応の識別番号による検索、出発物、生成物の全構造検索、部分構造検索の機能を備えている。反応中心の自動認識機能では、出発物、生成物を入力すると、出発物、生成物の構造間の最大共通骨格構造を認識し、これを出発物、生成物の構造と比較することによって反応中心、反応部位を自動認識した。このプログラムは、複数の化合物で構成される反応をも取り扱える、出発物、生成物の原子間の1対1の対応関係を得ることができる、また、複数の反応中心候補が得られる場合には、その可能性をすべて提示するという特徴を持つ。反応経路の列挙機能では、反応中心の自動認識から得られた結果を基に、出発物中の切断された結合を、順番に切断し、その切断点を組み上げるという操作を繰り返すことで、反応の経路を列挙する。また、この過程では、反応によって得られた反応生成物をも考慮することができる。この反応経路列挙のプログラムは、新たな経路の発見にもつながることが充分予想されうる。反応における知識の自動獲得においては、事実的知識から、単位知識を構築し、それをさらに体系化し知識ベースを構築していくという方法論を提案し、それに従って、有機合成化学に必要な単位知識の獲得とその利用に関する研究を行った。先に述べた反応データベース管理システムにより事実的知識を構築し、反応中心認識などにより得られた個別知識とを融合させ、そこから二重結合生成に関する単位知識の獲得を行った。そして、獲得した単位知識を利用する機能を開発し、さらに知識を拡張(学習)していく手法につ

いても考察した。

以上の研究で、有機化学に関する知識を自動的に獲得することが可能になった。ここで得られた知識は、有機合成設計システムにおいて活用できるであろう。また、本研究での一連の知識処理の方法は、化学におけるエキスパート・システム構築に必要な知識の獲得の一つの方法として有用なものである。