

非線形光学材料を対象とした自己組織型情報ベースシステム

○ 宇陀則彦[†]藤原 譲[‡]

Information-Base Systems for Nonlinear Optical Materials

Norihiko Uda[†]Yuzuru Fujiwara[‡]

Research and development of material design require advanced systems which have sophisticated functions such as analogical reasoning, inductive inference, abductive inference as well as retrieval and deductive inference.

This paper describes Information-Base Systems with Self Organizing Receptor Interconnections(IBM/SORITES) for nonlinear optical materials design. The systems store full-text, and structuralized information including semantic relationships which are extracted from full-text automatically, and have capabilities of navigation in the conceptual space as well as in the logical space and analogical reasoning in addition to retrieval and browsing.

1 はじめに

材料開発などの先端科学技術分野における情報システムには、研究開発を支援するために、検索、演繹推論だけではなく、類推、帰納推論、仮説推論などの高度な機能が必要である。従来のデータベース、知識ベースのほとんどは、表現の制約、管理の制約のために研究開発に必要な情報を格納することが困難である。

本稿は、非線形光学材料情報を例にとり、研究開発用自己組織型情報ベースシステムの概要について述べる。本システムは、フルテキスト、意味的關係に対応した構造化情報を格納し、検索、ブラウジング、概念空間や論理空間におけるナビゲーション、類推などの機能を持つ。

2 自己組織型情報ベースシステム

自己組織型情報ベースシステムは、従来のデータベースシステム、知識ベースシステムの持つ情報と機能を越える新しいタイプのシステムである。情報の自己組織化とは、情報自身に基づいて原情報を自動的に構造化することを指す。計算機に格納される情報は、非常に大量で複雑であるため、人手による処理は、困難であり、自己組織化は必須である。

[†]図書館情報大学

[†]University of Library and Information Science

[‡]筑波大学 電子・情報工学系

[‡]Institute of Information Science and Electronics, University of Tsukuba

図1は、自己組織型情報ベースシステムの構成である。情報ベースシステムは、研究開発に必要な多種多様な情報をすべて格納する。すなわち、一次情報、二次情報、概念構造、論理構造である。概念構造、論理構造の意味的關係は自動的に抽出され、構造化される。構造化するための手法は、対訳辞書を利用して同値關係、階層關係を抽出する C-TRAN、造語規則を利用して階層關係、關係關係を抽出する SS-KWIC、構文パターンを利用して、様々な意味的關係を抽出する SS-SANS、既存のソースを統合する INTEGRAL などである。また、格納した情報に対して検索、演繹推論だけではなく、概念構造、論理構造を基にした類推、帰納推論、仮説推論などを行なう。

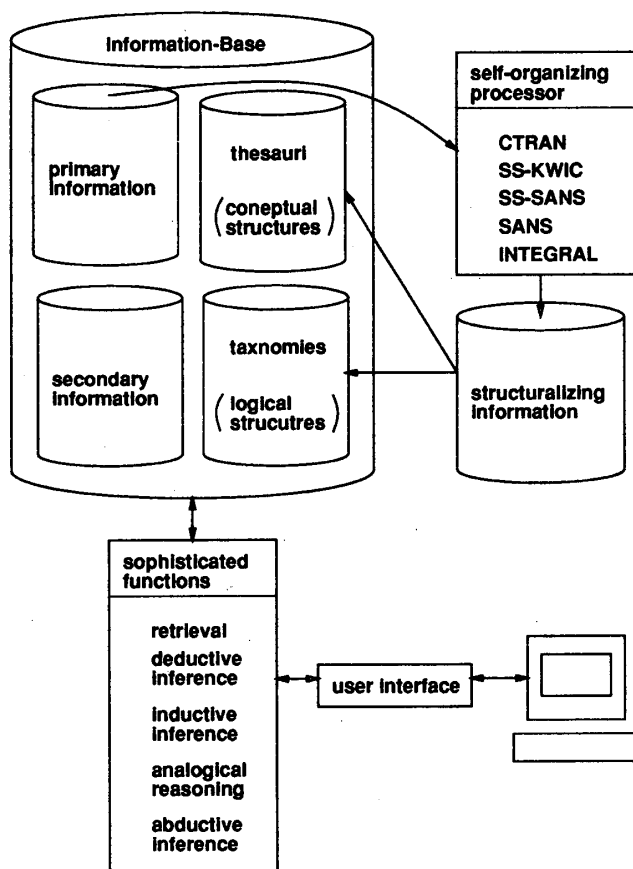


図1: Architecture of IBS/SORITES

システムは、人間のよう飛躍的な推論はできないが、思考の偏りがないため、場合をつくして、全ての可能性を考慮し、結果を出力する。研究者は、システムから得られた結果を考慮し、研究開発に役立てることができる。研究者が納得するような妥当な推論を行なうためには、対象となる分野の意味的な關係を抽出し、現実の情報に即した形で表現し、構造化するのが効果的である。

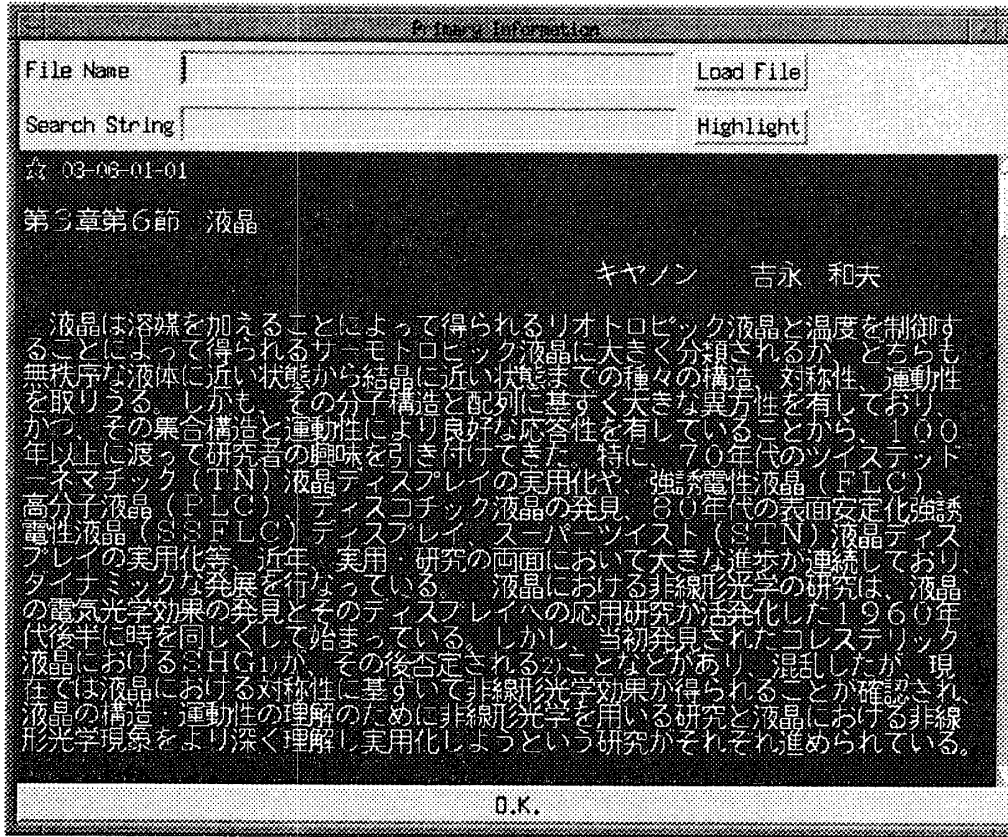


図 2: The Text Window of IBS/SORITES

3 非線形光学材料情報への応用

本稿では、非線形光学材料を例にとった情報ベースシステムについて述べる。非線形光学材料は、ハンドブック形式のフルテキストを基本とした情報である。

図 2 は、ある章のフルテキストである。テキスト上には、色のついた用語がある。この用語に対してもマウスのポインタが反応し、左ボタンをクリックすることにより、図 3 のように SS-KWIC による階層構造が表示され、右ボタンをクリックすることにより、図 4 と 図 5 に示すように、クリックした用語に対する概念構造と論理構造が表示される。

図 3 は、造語規則を利用した SS-KWIC による概念のナビゲーションの例である。用語をクリックすると、階層関係にしたがって用語をたどることができる。ある用語に修飾語がついている場合、修飾語がついている用語をその用語の下位概念とみなす。ここでは、電気光学結晶、無機光学結晶、光伝導性電気光学結晶は、光学結晶の下位概念であり、結晶は、上位概念である。また、造語規則は機械的に判断できるため、計算機による大量の処理が可能である。

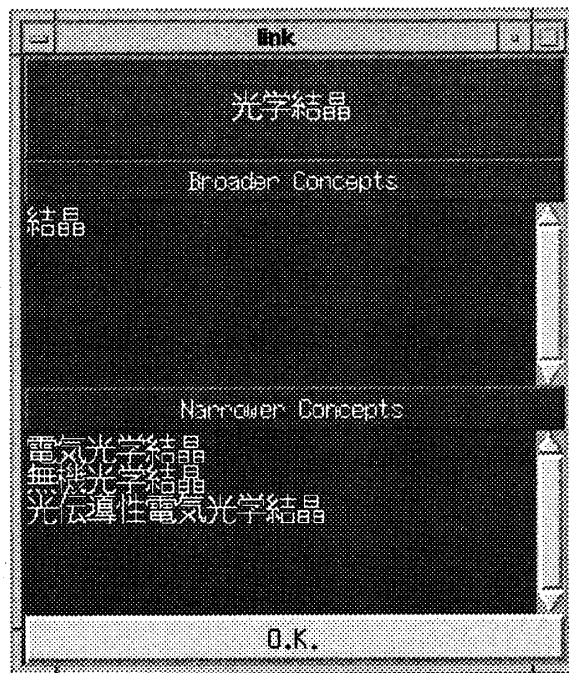


図 3: The SS-KWIC Window of IBS/SORITES

概念構造、論理構造上の用語に対しても、マウスのポインタが反応し、用語に対してナビゲーション、類推などを行なえる。ナビゲーションにおいては、どの意味的關係でナビゲーションを行なうかを操作作用のウインドウで選択できる。これにより、異なった観点から概念をたどることができる。

また、図 4 と 図 5 の概念構造、論理構造からラベルウインドウの類推のボタンをクリックすることにより、類推を行なうことができる。

図 6 は、類推結果の例である。図 5 の論理構造では、PHNMA が持つ性質が非線形光学効果を引き起こすことを示している。一方、図 4 の概念構造では、PHNMA と PPNMA がともに結晶高分子であること、すなわち両方の性質は類似していることを示している。それゆえ、PPNMA の性質も非線形光学効果を示すのではないかという類推を行なった。

図 7 は、構造化のためのファイル構成を示している。原情報とインデックスファイルから、用語のウインドウ上での位置情報を示すタグファイルを構築する。またインデックスファイル内の用語を SS-KWIC の手法により、下位語ファイルを構築し、それをインバーテッドすることにより、上位語ファイルを構築する。これらを概念構造ファイルにマージする。

図 8 に、本システムの全体の流れを示す。システムを起動すると、初期画面が表示され、そこで光学材料情報を選択する。すると、目次画面が現れ、章を選択すると、フルテ

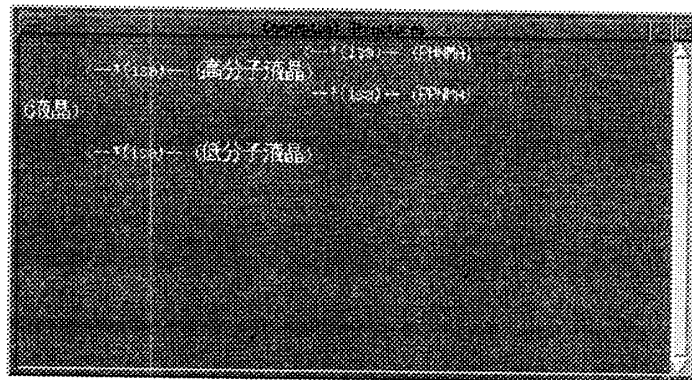


図 4: The Conceptual Structures Window of IBS/SORITES

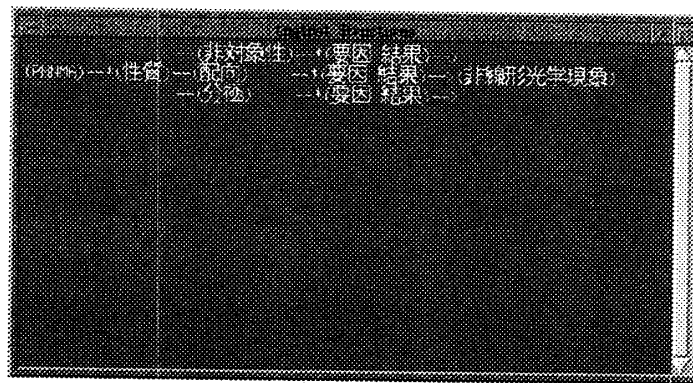


図 5: The Logical Structures Window of IBS/SORITES

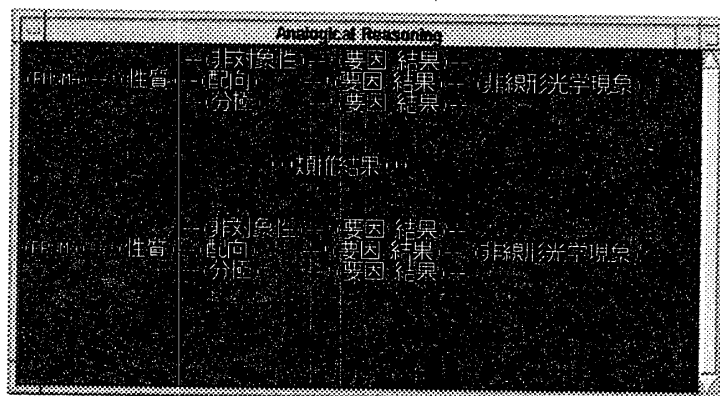


図 6: The Window of Analogical Reasoning

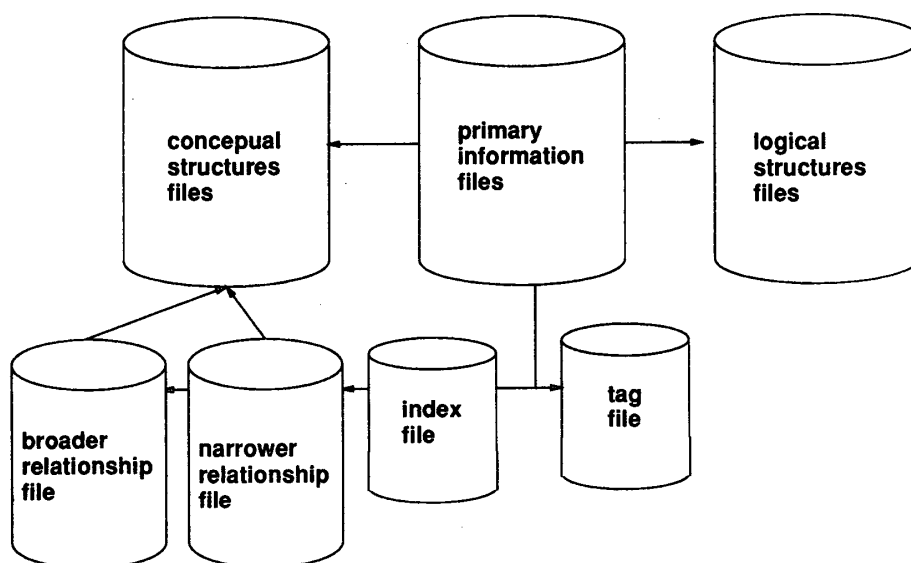


図 7: Files of Structural Information

キストウインドウが表示される。テキストウインドウ上の用語をクリックすると、その概念構造と論理構造が表示される。これら概念構造、論理構造に対してナビゲーション、類推の処理を行なう。

4 考察

研究開発支援システムを構築する場合、研究開発情報特有の問題を考慮する必要がある。研究開発の途中の段階で様々な結果がでるが、その時点では、真か偽か確定していない情報が多い。従来のシステムでは、これらどちらともとれる情報を格納し、処理することは困難である。また、扱う分野によって情報の形式は異なる。従来のシステムでこれら进行处理しようとする、表現形式が固定的であり、限定されているため、本来持っている情報の内容を歪めてしまったり、入力することができない場合がある。また、管理の制約によって同義性、多義性を正確に処理できない場合がある。さらに、研究をすすめていく段階で未知の概念が判明することもしばしばである。

このことは、結局、情報の本質的性質と密接な関連があり、情報の意味処理は避けられない問題である。自己組織型情報ベースシステムにおいては、意味を構造として処理する。すなわち、現実の情報が持つ構造をそのまま計算機の内部に再構成して組織化する。構造がそのまま計算機に組織化されれば、属性の継承、類似関係、各種の推論などの意味処理は比較的容易である。従来のシステムでは、現実の情報の性質を反映していないデータに対して処理を行ってきたが、単純なデータに対して複雑な処理を行っても妥当な結果を出すのは困難である。適切な意味処理を行なうためには、情報の構造を構築するの

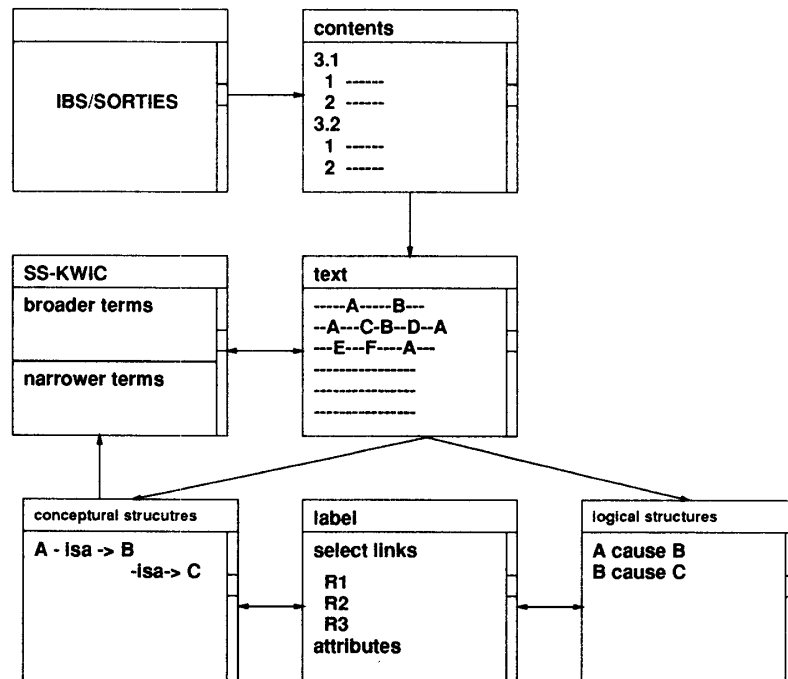


図 8: Windows of IBS/SORITES

が早道である。

次に自己組織化であるが、今回、そのシステム化の一環として SS-KWIC のインプリメントを行なった。用語間の階層関係、関連関係を造語規則にしたがって自動的に抽出した。SS-KWIC は、大量の情報を機械的に組織化できることが特徴である。もちろん、階層関係、関連関係は、造語規則によらないものもあるが、これは、別の手法で抽出する。まずは、造語規則によっておおまかな構造を構築する。特に、専門用語の場合、比較的造語規則による階層関係は多く、SS-KWIC は有効である。自己組織化とは、一つの手法で全ての構造を完全に構築するのではなく、複数の手法を用い、自身の結果を相互に利用し徐々に構造を構築する。

今後は、他の構造化手法をインプリメントし、自己組織型情報ベースをシステム化していく。

5 結論

本稿は、非線形光学材料情報を対象とした自己組織型情報ベースシステムについて述べた。本システムは、フルテキスト、構造化情報を格納し、その情報に対してナビゲーションおよび類推を行なった。また、造語規則を利用して自動的に階層関係を抽出した。このシステムにより、研究開発を支援するシステムの方向性を示した。

参考文献

- [1] Y. Fujiwara, N. Uda, and X. Zhang. "Analogical Reasoning in Polymer Information-Base Systems". *Proceedings of 13th International CODATA Conference*, Beijing, October 1992.
- [2] Y. Fujiwara, Z. Q. Wang, S. Q. Zheng, and Y. Q. Luan. "The Multicategorical Structures of Information for Inferences and Reasoning in the Self-organizing Information-Base Systems". *Proceedings of CAMSE'92*, 1992.
- [3] P. H. Winston. "Learning and Reasoning by Analogy". *Communication of the ACM*, 23(12):689-703, 1980.
- [4] M. H. Berstein. Combining analogies in mental model. "*Analogical Reasoning*". Kluwer Academic Publishers, 1988.
- [5] D. Gentner. The mechanisms of analogical reasoning. "*Similarity and Analogical Reasoning*". Cambridge University Press, 1989.
- [6] S. Kurosaki, M. Nagao, S. Sato, and M. Murakami. "A Method of Automatic Hypertext Construction from an Encyclopedic Dictionary of a Specific Field". *Artificial Intelligence*, 7(2):336-345, April 1991. (in Japanese).
- [7] T. H. Nelson. "The Hypertext". *Proceedings of the World Documentation Federation*, 1965.
- [8] F. G. Halasz, T. P. Moran, and R. H. Trigg. "NoteCards in A Nutshell". *Proceedings of CHI + GI*, 1987.