

化学情報サービスに向けたナレッジグラフ構築Ⅱ

高分子物性の理解

Knowledge Graph Construction for Chemical Data Utilization Service II



株式会社富士通研究所

池田 紀子

人工知能研究所 デジタルナレッジプロジェクト、技術士（応用理学／総合技術監理部門）、電子・光学デバイス材料の設計および分析、並列処理および分子モデリングの研究、特許読解支援システムの開発、ナレッジグラフの研究に従事。

✉ nona@fujitsu.com

TEL 044-754-2652



株式会社富士通研究所

田中 一成

人工知能研究所 デジタルナレッジプロジェクト、テキストマイニング技術の研究、特許読解支援システムの開発、ナレッジグラフの研究に従事。

✉ tanaka.kazunari@fujitsu.com

TEL 044-754-2652



株式会社富士通研究所

吉田 宏章

人工知能研究所 デジタルナレッジプロジェクト、環境・エネルギー材料の開発、エネルギーマネジメント、特許読解支援システム、ナレッジグラフの研究に従事。

✉ yoshida.hiro-15@fujitsu.com

TEL 044-754-2652

1 はじめに

化学分野の特許（以下、化学特許）文書には、他分野とは異なる請求項の書き方や化学物質名の膨大な羅列がある。これらのことが原因で、人が化学特許を読み、理解し知識を得、さらにイノベーション創出に活かすためには、予想以上に高度なスキルと多大な労力が必要である。例えば、イノベーションの発生過程は、「研究」→「開発」→「設計」→「製造」→「販売」のような直線的な流れでなく、各々が連鎖的に関連し、分野間でフィードバックが起こりつつ発生する連鎖モデルである^[1]。各分野間では、それぞれの専門知識の間に大きなギャップが存在するため、分野をまたぐ材料の設計やイノベーション創出には、専門知識同士が紐づき統合されたナレッジ基盤が必要である。そこで筆者らは、化学特許の

読解支援に活用できる化学ナレッジ基盤の開発を進めている。これまで、言語的観点に化学的観点を加えた認識・解析モデルアプローチで化学特許の読解支援に取り組んできた。前報の「化学情報サービスに向けたナレッジグラフ構築」では、化合物や製造方法についての知識をナレッジグラフとして連携させることで、知識の利活用拡大と調査の負担削減の可能性が高まるとした^[2]。ナレッジグラフとは、グラフ形式で表現された知識ベースで、効率的に知識を表現、統合、解析することができるという特徴をもつ。今回、未着手であった物性の読解支援についてナレッジグラフを用いて検討したので、報告する。第2節では、化学ナレッジ基盤と読解支援の基本的な考え方、第3節では、化学情報のアプリケーションの1つとして、検索・読解支援システム、第4節では、物性情報の抽出について説明する。

2 化学ナレッジ基盤の活用

クライン (Stephen J. Kline) の連鎖モデルをもとに科学技術振興機構研究開発戦略センターが作成したイノベーションモデル^[1]を参考に、アイデアの創出、知識の流通・活用を目指す化学ナレッジ基盤の位置づけを図1に示す。化学ナレッジ基盤を活用して、単なる一組の相関だけではなく、複数の関係性の特徴を加味することで、材料開発や製品開発の信頼性向上を目指している。

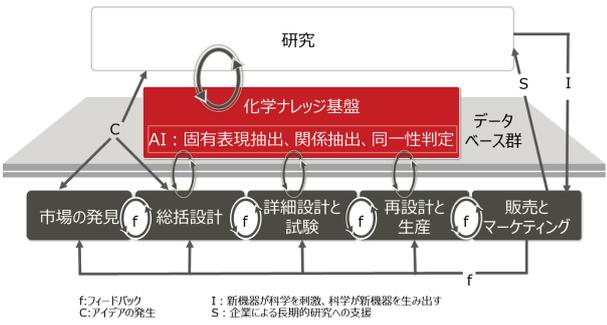


図1 イノベーションモデルと化学ナレッジ基盤

2.1 化学特許の読解支援

図2に示すように、化学ナレッジ基盤を活用した化学特許読解支援の主な要素技術には、文書分類、化合物や物性等の固有表現抽出、関係抽出、化合物の同一性判定などがある。



図2 化学ナレッジ基盤の主な要素技術

これまで、特許文書の一文中から化学物質や化合物群を抽出でき、文同士の関係性から製造方法を捉えることができた。図3に示すように、特許文書から化学に関する用語や知識を自動抽出し、体系化するための方法を考案している。それに基づいて、以下4つの内容で化

- 学情報検索・読解支援システムの開発を行ってきた。
- ①特許文書をコーパスとして用いることで、有機化合物の化学物質名と化学式の対応関係を抽出^[3]
- ②化学物質情報の知識化について、特許読解支援システムの実用化とオープンデータとの連携^[4]
- ③オープンデータによる体系名や別称の同一性判定のために日化辞データを利用した言い換えルール辞書を構築^[5]
- ④化合物のナレッジグラフや製造方法情報のナレッジグラフ構築の試み^[6]



図3 化学特許読解支援のアプローチ

2.2 化学特許の物性

物性とは物質の力学的、熱的、電気的、光学的性質であり、モノづくりの根幹となる知識である。材料の物性と性質を紐づけることが今回の課題である。化学特許の物性に関する記載には、以下の問題がある。

- ①物性の定義、測定方法、計算式などの明確な記述が不足
- ②物性の技術的意義の明確な記述が不足
- ③実施例のデータやパラメータが不足
- ④余分な記載

特許文書中の記載に上記のような不備がある場合、物性を理解することが困難である。そこで、特許文書だけではなく、他の文献との連携が必要であると考え、物性に関する複雑な関係について、ナレッジグラフを用いた抽出情報の拡充に向けて漏れなく必要な情報を記述していくことを試みた。

2.3 高分子化学特許の読解支援

化学分野の中でも高分子を例にあげる。材料開発は材料メーカーだけに留まらず、製品メーカーの関与も大きい。たとえば、材料を採用する製品設計者が、プラスチック成形品の必要な物性をきちんと規定することや開発過程でできた候補材料の実用的な観点の評価試験を行っている。化学ナレッジ基盤の使用者の一例として、プラスチック材料開発の材料設計者や製品開発の製品設計者を

想定する。両設計者は新材料や新製品の先行特許調査を行いながら、技術を見極め、物性と評価試験情報を深く理解しなければならない。この業務は研究部門や知財部門とも連携をとりながら進めていくので、化学ナレッジ基盤は関係部門に有効なシステムでなければならない。本報では、プラスチックの性質と物性の関係のナレッジ化例について述べる。

また、プラスチック製品強度設計には、以下の課題があるとされている^[3]。

- ①設計に関するデータベースの蓄積が比較的少ない
- ②粘弾性特性を有するため、材料力学を適用できる設計範囲が比較的狭い
- ③絶対強度が低いので、設計の許容値（応力、ひずみ）を狭くした範囲で設計をせざるを得ない
- ④使用条件によって性能が変化しやすい
- ⑤製品形状、成形条件、二次加工などの要因が関係するので、強度に影響するパラメータが多い

3 化学ナレッジ基盤システムとツール

3.1 化学情報検索・読解支援システム

化学情報の活用支援のために、検索・読解支援システムを試作し、そのシステム構成を図4に示す。

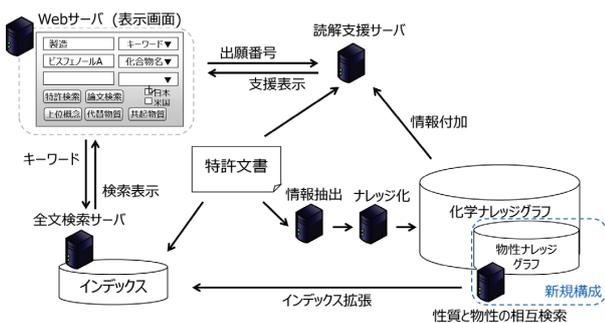


図4 化学情報検索・読解支援システムの構成

3.2 化学特許検索ツール

化合物名や物性情報を抽出し、関係抽出、同一性判定を可能にした試作ツールを紹介する。キーワード、書誌情報、化合物名（別称に対応）による検索を行える入力画面を図5に示す。検索結果の特許明細書について、化合物名や物性情報を自動抽出してハイライトした出力画面の一部を図6に示す



図5 検索入力画面

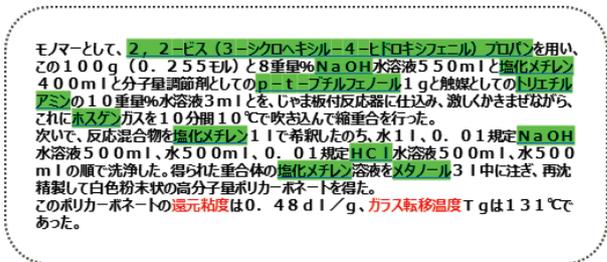


図6 検索結果の特許明細書

4 物性情報抽出

4.1 材料の性質と物性の関係

化学特許文書中には、図7に示すように、産業上の利用分野の中にポリカーボネート系樹脂組成物（以下、PC）などの材料の性質が記載されることがある。図7の例では、PC材料の性質として、明示的に「材料として好適な」「などに優れる」等と示されている。

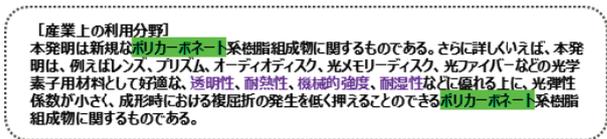


図7 特許明細書からの抜粋例：産業上の利用分野

図8にPC材料の優れた性質と物性を示す。図7から表1に示す4つの優れた性質を抽出できる。図8の表2に前報で抽出した、最終生成物の物性を示す[2]。ここで、表1が示す優れた性質群と表2が示す物性群を抽出することができるが、性質と物性の各々の関係、すなわち、発明の具体的な数値が示す効果を理解することが困難である。

表1 優れた性質

表2 物性

優れた性質
透明性
耐熱性
機械的強度
耐湿性



物性	値
還元粘度	0.48dl/g
ガラス転移温度	131℃

図8 PC材料の優れた性質と物性

そこで、計算機上に表現したい優れた性質と物性の関係のナレッジ化を図9に示す^[7-9]。図中a図は還元粘度、b図はガラス転移温度の優れた性質との関係を示す。一般に、重合で得られる最終生成物は、比較的簡単な構造単位が多数繰り返してつながった巨大な分子である。高分子の性質は個々の高分子鎖の形やサイズに著しく依存する。高分子を特徴づける基本的な要素は構造と分子量である。このPC最終生成物の優れた性質について、「還元粘度」から機械的強度を、「ガラス転移温度」から耐熱性を把握できる。これらの物性からPCの優れた性質を把握するプロセスについて述べる。

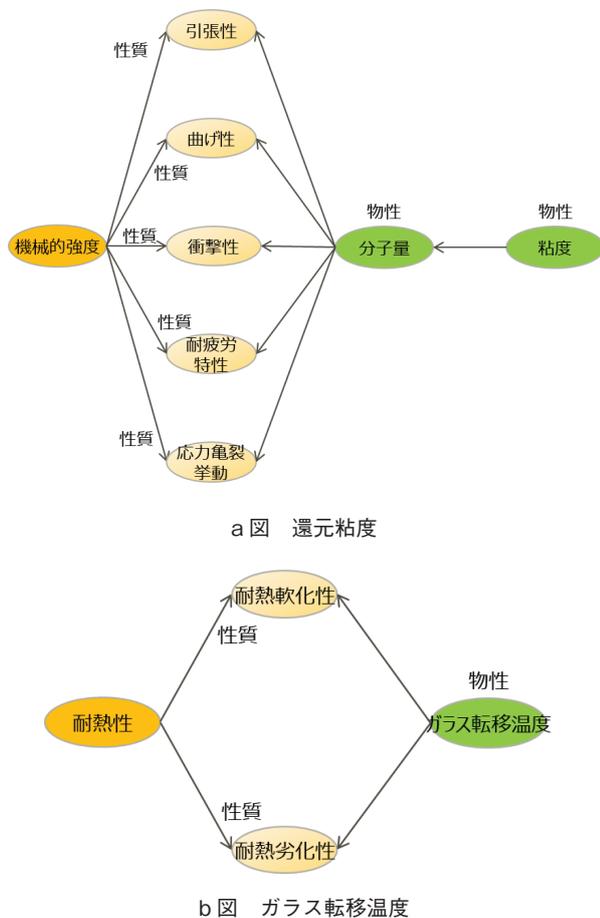


図9 優れた性質と物性の関係のナレッジ化

4.2 材料の物性の関係具体化

還元粘度と測定の関係のナレッジ化を図10に示す^[7,9,10]。還元粘度は測定から求めるが、測定方法と測定条件が複数あるので、その関係知識を抽出することが重要である。ナレッジグラフを用いて、それらの違いを確認できるようにするコンセプトである。

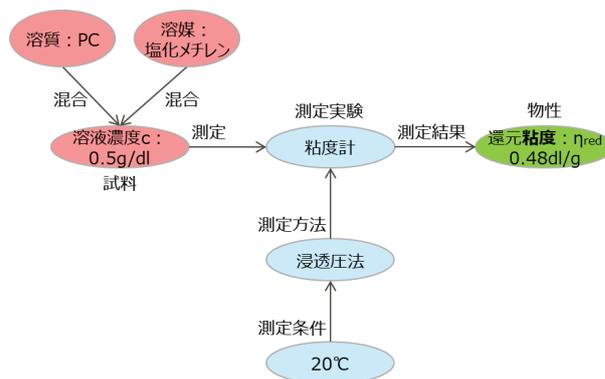
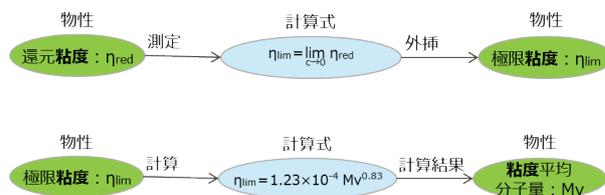


図10 還元粘度と測定の関係のナレッジ化

還元粘度と分子量の関係のナレッジ化を図11に示す^[9,11]。測定で求めた還元粘度から極限粘度を外挿し、得られた極限粘度を代入した計算式から粘度平均分子量を求める2段階の方法である。粘度平均分子量を求める計算式について測定方法と測定条件の溶媒に関する報文が2桁以上あるので、その関係知識を抽出することが重要である。



上図は極限粘度、下図は粘度平均分子量

図11 還元粘度と分子量の関係のナレッジ化

図12に耐熱性と物性の関係のナレッジ化を示す^[8-10]。図中上部は短期性質、下部は長期性質で時間変化を示す。両性質は次に示す構造変化を伴う。熱軟化とはガラス転移温度 T_g を超えるとPC高分子鎖が動き始め、柔らかくなり変形しやすくなる。熱劣化とは熱により樹脂の化学結合がラジカル的に切れていく分解である。長期間、大気中の酸素に晒され、化学結合が切れやすくなってしまふ酸化でもある。高分子の主鎖が切れていくと分子量が小さくなりネチャネチャした触感に変わっていく。よって、長期性質に機械的強度の物性である分子量が出現している。ナレッジグラフを用いて、分子量の関係知識を引用できるようにするコンセプトである。

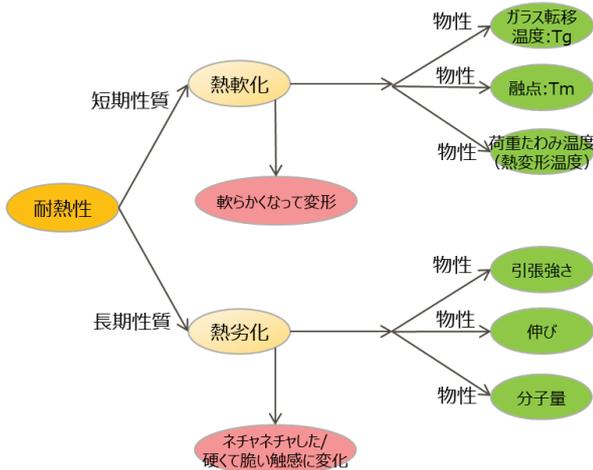


図 12 耐熱性と物性の関係のナレッジ化

図 13 に熱劣化の機構と計算式の関係のナレッジ化を示す^[9,10,12]。種々の温度領域や環境で変化の様子が異なるので、その関係知識を抽出することが重要である。

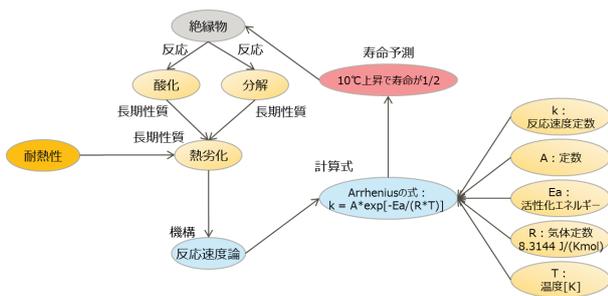


図 13 熱劣化の機構と計算式の関係のナレッジ化

図 9 ~ 13 に材料の性質と物性の関係の一例を示したが、これらの関係図をリアルタイムに表示していくことは、たいへん困難である。そこで、その仕組み構築に向けた取組みとして、有効なナレッジツール群の提供を目指している。そのため、物性名、対象、単位、取得方法（測定方法か計算式）、取得に必要なパラメータを参照するためのデータを収集中である。また、今回は材料の観点であったが、今後、製品の観点で「性能」や「機能」を追加していく。

5 まとめ

物質の化学的屬性（物性）や化学反応性に注目した原子団や化合物の全体構造を把握することや、静的、動的、潜在的な分子間相互作用を把握できるようにしていくことで、新規化合物や新しい製造方法にたどり着く支援ができるようになる。ナレッジグラフ等を用いて効率的な集合知の利用方法を確立して、構造や分子間相互作用の探索や推薦を可能にしたい。しかし、情報の有効性や信頼性の判断基準をどう構築するかが大きな課題である。そのためにも化学知識抽出システム開発のために必要なコーパスやオントロジー等を整備しなければならない。これらは、実際に化学知識の抽出を行ったり、その結果を評価したりする際にも使用していく。複雑な化学知識をその本質を損なうことなく計算機上に表現し、化学知識を利用する技術を開発しなければならない。そのために化学オントロジーの表現方法や利用方法などを開発することも必要である。

現場では、当初の想定が外れてしまうことがあると踏まえ、知識を段階的に獲得していくことを具体化したい。それにより、その知識に応じて以降の設計などを修正していく場合に利用できると考える。さらに、知識の深掘り（当初には持っていなかった知識獲得）による構造や分子間相互作用の理解と類似から求めるアナロジー（類推）も目指していきたい。

参考文献

- [1] 平成18年版 科学技術白書 コラム07 イノベーションとは (2006)
http://www.mext.go.jp/b_menu/hakusho/html/hpaa200601/column/007.htm
- [2] 田中一成, 池田紀子: 化学情報サービスに向けたナレッジグラフ構築, Japio YEAR BOOK 2018, p.256-262 (2018) http://www.japio.or.jp/00yearbook/files/2018book/18_3_07.pdf
- [3] 池田紀子, 田中一成: 特許文書からの化学物質情報の抽出, Japio YEAR BOOK 2015, p.274-281 (2015)
http://www.japio.or.jp/00yearbook/files/2015book/15_3_06.pdf
- [4] 池田紀子, 田中一成: 特許文書から抽出した化学物質情報の知識化, Japio YEAR BOOK 2016, p.204-209 (2016)
http://www.japio.or.jp/00yearbook/files/2016book/16_3_06.pdf
- [5] 田中一成, 池田紀子: オープンデータを用いた化学特許情報活用へのアプローチ, Japio YEAR BOOK 2017, p.206-211 (2017)
http://www.japio.or.jp/00yearbook/files/2017book/17_2_11.pdf
- [6] 課題解決に導く プラスチック製品の強度設計&トラブル防止策
<https://tech.nikkeibp.co.jp/dm/atcl/seminar/16/010600295>
- [7] JIS K 7367 (ISO 1628) プラスチック-毛細管形粘度計を用いたポリマー希釈溶液の粘度の求め方
- [8] JIS K 7121 プラスチックの転移温度測定方法
- [9] 高分子データベース (PoLyInfo) - 物性
<https://polymer.nims.go.jp/PoLyInfo/guide/jp/Property.html>
- [10] ユーピロン物性編
https://www.m-ep.co.jp/product/brand/iupi_nova/bussei.html
- [11] 極限粘度
https://www.ube-ind.co.jp/usal/documents/b112_141.htm
- [12] 電気用品安全法
<https://www.meti.go.jp/policy/consumer/seian/denan/>