

サイバーデザインによる特許データベース有効利用

米澤 幸子

法政大学大学院 IT Professional Course

〒102-8160 東京都千代田区富士見町 2-17-1 ITPC 事務室

it013343@itpc.i.hosei.ac.jp

國井 利泰

法政大学大学院 IT Professional Course

〒102-8160 東京都千代田区富士見町 2-17-1 ITPC 事務室

tosi@kunii.com; <http://www.kunii.com/>

要旨

本論文は、セルモデルの応用として、WWW サーバ上のデータベースに蓄積された情報から、利用者の目的に沿った特許情報を効率よく検索する方法を考察する。

On effective utilization of patent databases by cyber design

Yukiko Yonezawa

Hosei University IT Professional Course

Tosiyasu Kunii

Hosei University IT Professional Course

Abstract

The cell model, is applied to search patent information effectively to meet the requirements of given applications on the web.

1. 知的所有権の保護の変遷

従来、知的所有権の保護の対象は、新規な技術に対するものであり、機械、電気、電子等の工業分野に特化したものと考えられてきた。従って、先行技術調査、関連技術調査の対象もその分野に特定され、同業者の動向調査、あるいは刊行物などの限定された資料を検索することにより、ある程度の情報収集が可能であった。

しかし、近年、技術の複合化がすすみ、前述し

たような限定された資料の調査のみでは、既存の権利に抵触する恐れが出てきた。特に金融関係では、サービスをインターネット上で行い、取引にネットワーク、コンピュータ関連の技術を利用するほか、金融取引というサービスについても多様化している。さらに、ビジネスモデルという各社独自のサービスが展開され、それに新規技術が複合されると特許取得が可能になるなど、今後、他人の権利の動向に注目せざるを得ない環境となってきている。

そこで、情報結合モデルであるセルモデルを応用し、複数分野にまたがる技術調査を可能とし、相互に関連する技術を効率よく検索すること、またその結果からデータベースを作成し、更新、変更を可能にすることを、本論文の目的とする。

2. 技術調査方法の変遷

従来の技術調査の方法としては、刊行物によるものが一般的であった。また特許庁による新規性調査の対象も、配布され、しかも発行年月日が明確である刊行物や、学術論文の要旨などが用いられることが多かった。これは、特許庁の審査官が入手可能な手段であり、また出願された発明について権利化を認めるか否かの判断材料として有用であるという理由による。

しかし、近年、インターネットが爆発的に普及し、その情報源としての価値が無視できなくなっている。特に、刊行物と比較して、低コスト、少労力で研究成果の発表が可能であることから、各企業が独自で自社技術を掲載する事例も増加している。また、近年、各サイトへのアクセス数が増加し、その情報の証拠能力としての適格性も認められるようになってきていることから、新規性調査の対象としても用いられるようになってきている。

そのような状況において、先行技術調査や他人の権利の動向調査に、インターネットを情報源として利用し、さらに、そのデータをもとにしてデータベースを構築することへの需要が、高まっている。

また、従来のデータベースはリレーショナルデータベースが多く使用されるが、このデータベースはデータの属性とそれらの関係性があらかじめ定められており、後の変更作業が困難であることから、これらの作業を容易にするような新たなモデルのデータベースの作成が求められている。

3. セルモデルを応用したデータベース構築

セルラーモデルとは、著者の一人である國井によって1999年に発表されたモデルで、現実のローカルな情報とグローバルな情報の結合が可能で、現実世界・サイバー世界のすべての事象を射影しうる情報空間構築モデルである。

(1) セルモデル

サイバーワールドモデルとして、“セル構造空間レベル”は、セル空間構造に基づいているが、そのようなサイバーワールド空間は、概念的及びデータのモデリングに共通するグラフ理論レベルに基づいているものより一般的な礎となり、情報モデルを、境界を持つあるいは持たないセルとして、認識により及び計算によりオブジェクトを特定することができる。境界を持つセルは“閉じて”おり、境界のないセルは“開いて”いる。ここで、 n 次元のセル、 n セルは、整数である n 次元の球と、トポロジー等価空間である。開いた n セルを e^n と表し、閉じた n セルを B^n と表す。 B^n は、整数

$$\text{数 } \text{Int } B^n = \overset{\circ}{B}^n, \text{ と表し、式 } \partial B^n = B^n - \overset{\circ}{B}^n = S^{n-1}$$

は、 B^n の境界であり、それは $(n-1)$ 次元である。セルモデルにより、普遍量としてセル次元及び接合性を維持することで、セル結合及びセル分解ができる。オブジェクト識別は、識別マッピング(しばしば商マッピングと呼ばれる)によりシステムティックにできる。ここでは、次元は自由度を意味する。後に、われわれは、このデータベーススキーマ結合、およびスキーマ分解は、セル結合及びセル分解の特別なケースであることを示す。

次元の例を挙げよう。例えば、サイバーワールドでは、属性を持たないオブジェクトは、属性値を変える自由度を持たず、それゆえ、サイバーワ

ワールドのオブジェクトにおける次元はゼロとなる。われわれは、これを、プレゼンテーションレベルで実現する。属性は、特定の性質あるいはオブジェクトに固有する特性の、相互に独立した集合である。一つの属性を持ち、その属性の値を変えることができるオブジェクトは、線として表現することができる。同様に、2及び3の自由度を持つ、属性が2及び3のオブジェクトは、これを、平面及び球として表現することができる。n個の自由度をもち属性がnであり、それゆえ、その次元がnであるオブジェクトは、これを、n次元の球として表現することができる。リレーショナルモデルは、リレーショナルスキーマとしてn属性をもつオブジェクトとして表現でき、n列のテーブルとして具体化できる。

接合性は、アタッチングマップと呼ばれる、連続性と全射マッピングにより定義される。関数fにより $X \rightarrow Y$ が全射であるマッピングとは、

$$(\forall y \in Y) (\exists x \in X) [f(x) = y]$$

ということの意味する。

関数fにより $X \rightarrow Y$ が連続であるマッピングとは、Yの部分集合であるAが、 $\{f^{-1}(y) | y \in A\}$ がXにおいて開いており及びそのときのみ、Yにおいて開いているということの意味する。

$$Y \sqcup_f X = Y \sqcup X / \sim$$

トポロジー空間であるX、及びYの排他和は、アタッチングにより得られるアタッチング空間である。 \sqcup は、排他和を示し、しばしば+記号で表記される。 \sim は等価関係である。等価関係は、単に、シメトリック、リフレクティブ、トランジティブの関係である。それは、集合理論等価関係であってもよいし、トポロジー的等価関係であってもよいし、地理的等価関係でもあってもよいし、あるいは、ホモトピー的等価関係であってもよい。そのトランジティブ性は、空間を等価クラ

スと呼ばれる、部分空間の排他和に分割する。

ここで、アタッチングマップの一般的な定義を記述する。全ての等価クラスの集合は、 X / \sim として知られており、Xの等化空間と呼ばれる。

$$X / \sim = \{x / \sim \in 2^X | x \in X\} \subseteq 2^X$$

アタッチングマップfは、全射及び連続的マップであり、

$$f: X_0 \rightarrow Y, \quad \text{where } X_0 \subset X$$

$X \sqcup Y / \sim$ は、等化空間であり、

$$X \sqcup Y / \sim = X \sqcup Y / (x \sim f(x) | \forall x \in X_0) = X \sqcup_f Y$$

これは、ウェブに情報を埋め込む際に、情報スキーマ統合及び情報統合のために後に使用する、特別なケースである。 S^{n-1} を閉じたnセル B^n の境界すなわち ∂B^n とする。そこでは以下の式が成り立つ。

$$S^{n-1} = \partial B^n = B^n - \text{Int } B^n = B^n - e^n$$

アタッチングマップfをサージェクティブ及び連続的マップとすると、

$$f: S^{n-1} \rightarrow X$$

付加スペースYは、等化空間として以下のように定義される。

$$Y = X \sqcup_f B^n = X \sqcup B^n / \{(f(u) \sim u | u \in S^{n-1})\}.$$

2つのホモトピーマップfとgを

$$f, g: S^{n-1} \rightarrow X \text{ とすると、} X \sqcup_f B^n \text{ および } X \sqcup_g B^n$$

は同じホモトピーのタイプを持つ。

$$X \sqcup_f B^n \simeq X \sqcup_g B^n$$

サイバースペースをトポロジー空間とすると、XからわれわれはJ.H.C.whiteheadに従って、整数Zにより見出しづけられるXの部分集合、セル X^p の有限的または無限的連続を再帰的に構成することができる。

アプリケーションには、重要なセル空間がある。それらはサイバースペース複雑性及びマニフォー

ルドを含んでいる。

(2) ウェブ情報モデリング

ウェブ情報モデリングのためにわれわれがしなければならない最初のことは、どのようにサイバースペースが発生したか及びそれらが何であるのかを把握するため、ウェブと共有される情報世界として、サイバースペースの形態の本質を特徴付けることである。しばしば、サイバースペース X は、多くのウェブサイトでの局所的、多様な作成の結果としてウェブ上で創造されるケースが生ずる。会社の情報と異なり、われわれにスキーマの前提となる集合を提供してくれる情報管理者が存在すると仮定することはできない。情報の埋め込みを通して、われわれは局所的ウェブサイト X が何であるかを示す特別な情報を埋め込むことが可能である。もちろん、われわれは、勝手に情報の埋め込みを行うことはない。情報がウェブサイトでブラウズされた後は、埋め込まれている情報及びそれらを統合することによりウェブのサイトの情報に現れているわれわれが望んでいる情報を集める。これは、通常“設計による情報埋め込み”と一般的に呼ばれている情報の埋め込みである。というのは、埋め込まれる、あるいは埋め込まれないものに関する情報として適用される、規則的な集合が存在するからである。そのような統合ガイドは、何を及びどのように統合するかという設計ガイドとなる。

ウェブの情報埋め込みは、前述した Whitehead の具体化スキームにのっとり、局所的なウェブ世界から世界的なサイバースペースへと、行われる。どのように具体化された統合が行われ、 n 次元のサイバースペース X^n が得られるかを、より詳細に記述するため、ウェブの検索及び統合過程を説明する。

具体的統合は、2つのフェーズからなる。情報

スキーマ統合フェーズ及び情報統合フェーズである。第1のフェーズ、情報スキーマ統合フェーズは、次のように処理される。

1. 0次元のサイバースペース X^0 を創造するためウェブサイトの属性を持たない全てのセル B_i^0 を検索する。

$$X^0 = \{ B_1^0, B_2^0, B_3^0, \dots, B_j^0 \}$$

2. 1次元のサイバースペース X^1 を創造するためにウェブサイトの一つの属性を持つセル B_i^1 の全ての結合を検索し、それらの排他和 $\sqcup B_i^1 = B_1^1 \sqcup B_2^1 \sqcup B_3^1 \dots \sqcup B_k^1$ を、個々の境界要素を識別することにより、 X^0 へアタッチングマップ F を通して接合する。それにより、多様な1次元のサイバースペース X^1 を得ることができる。

$$X^1 = X^0 \sqcup F(\sqcup B_i^1) = X^0 \sqcup (\sqcup_i e_i^1)$$

where $i=1, \dots, k$,

アタッチングマップ $F: \sqcup_i \partial B_i^1 \rightarrow X^0$

3. 検索及び統合を繰り返すことにより、 $(n-1)$ 次元のサイバースペースを情報埋め込みを通して得ることができる。 X^{n-1} は $n-1$ 個の属性を持つ。 n 個の属性を持つ n 次元のサイバースペース X^n を統合するために、 $n-1$ 個の属性を持つ n セル B_i^n の全ての組み合わせを検索する。

それから、それらの排他和 $\sqcup B_i^n = B_1^n \sqcup B_2^n \sqcup B_3^n \dots \sqcup B_m^n$ を、 n セル B_i^n の n 個の属性のうち、それぞれの境界要素である $(n-1)$ 個の属性のそれぞれに一致させるアタッチングマップ G により、すでに作られた $(n-1)$ 次元のサイバースペース X^{n-1} にアタッチする。以上の処理を通して、多様な n 次元のサイバースペース X^n が得られる。

$$X^n = X^{n-1} \sqcup G(\sqcup B_i^n) = X^{n-1} \sqcup (\sqcup_i e_i^n)$$

where $i=1, \dots, k$, and アタッチングマップ $G: \sqcup_i$

$\partial B^{n_i} \rightarrow X^{n-1}$.

以上が、情報スキーマ統合の詳細である。

次に、情報統合フェーズであるが、これは、かなり単純である。これは、設計ガイドに基づいて、スキーマを統合する際、セルをアタッチするそれぞれのステップにおいて、全てのインスタンスを、セルアタッチメントにより作られるサイバースペースに含むべきか否かをチェックすることである。

前述した範囲においては、識別は、等価関係に基づく等価クラスによりなされる。セルモデルの本当の強さは、この面に見られる。ウェブ上ではかなり複雑で急速に変化するが、セルモデルの統合力により、真の理論的基盤とともに、ウェブ情報モデルを提供する。ウェブに基づく情報システムでは、設計ガイドはイントラネットのような局所的なものであるか、あるいは、ボーダレスのサイバースペースのような世界的なものであるかのどちらかである。設計ガイドは、セルに基づく情報システムの再利用可能なリソースである。

(3) 具体的事例

これから、ウェブ上の情報での情報構築及び検索について、ウェブ上での特許情報検索を通して、具体的事例について考察する。

① 情報スキーマ統合フェーズ

情報スキーマ統合フェーズにおいては、まずウェブから属性を持たないセル X^0 を検索する。

次に、1次元のサイバースペースを構築するために、ウェブサイトの一つの属性を持つセルの全ての結合を検索する。この属性には、検索の対象となる技術分野、たとえばネットワーク、コンピュータ、金融などの情報を用いる。それらの排他

和を、個々の境界要素を識別することにより、 X^0 へ接合する。

最後に、これらの検索及び統合を繰り返すことにより、 $(n-1)$ 次元のサイバースペースを、情報埋め込みを通して得ることができる。この繰り返行われる検索の属性には、検索の対象となる具体的なキーワードを用いる。たとえば関連技術の検索として、ネットワークの検索であれば、アプリケーション層、TCP/IP等の技術を実現する対象、金融の検索であれば、ネット取引、証券、株式取引などのサービス名等の情報を用いる。

それから、それらの排他和を、 B^n の n 個の属性のうち、それぞれの境界要素である $(n-1)$ 個の属性のそれぞれに一致させるアタッチングマップにより、すでに作られた $(n-1)$ 次元のサイバースペース X^{n-1} にアタッチする。以上の処理を通して、多様な n 次元のサイバースペース X^n が得られる。

このモデルの利点は、セルの検索と同値関係による接合により、求める情報の接合関係、接合空間を表現できることにあり、この点において、従来のデータベースと一線を画すものである。従来のデータベースにおいては、データを構築する際にスキーマを定義しなければならず、またそのスキーマの変更は、後に変更を必要とする場合においても再正規化を必要とし、容易ではない。また、従来のデータベース、リレーショナル・データベースにおいてはクエリを最適化しなければならず、また、XMLにおいては、スタイルシートを別途作成することにより、必要なスタイルの情報を入手することになる。

しかし、セルのデータベースにおいては、サイ

バーワールドを新たに構築するという概念を取り入れており、前述したような新たなスタイルの作成は必要とならない。しかも、種類の技術分野から特定の属性をもつデータを検索するのに適しているため、金融関係の業務をネットワークを通じて実現する等技術が複合化している今日においては、本データベースを活用するメリットは大きい。

②情報統合フェーズ

情報統合フェーズであるが、設計ガイドに基づいて、スキーマを統合する際、セルをアタッチするそれぞれのステップにおいて、全てのインスタンスを、セルアタッチメントにより作られるサイバーワールドに含むべきか否かをチェックすることにより行う。

前述した処理により得られたサイバーワールドは、(n-1)次元のサイバーワールドを、それぞれの境界要素である属性のそれぞれに一致させるアタッチングマップによりアタッチされたものであるため、求める技術分野の情報が膨大である場合、そのサイバーワールド構築にかなりの時間や労力を浪費する。しかし、情報統合フェーズでその絞込みを行えば、属性のそれぞれに優先順位をつけ、あるいは、必要な同値関係を設定することにより同値な属性を設定することにより、等化関係によりセル接合を行い、それにより情報量を軽減し、求めるサイバーワールドを容易に構築することが可能となる。また、実用上は応用に依存した処理になる。この場合、全体を再構築するのではなく、応用の範囲内の世界をアタッチングマップにより構成し、最適化することになる。

(i) ウェブ状態モデル

次に、既存のサイバーワールドから、新規のサイバーワールドを構築する場合を考える。この場

合、情報スキーマは、識別処理に従ったセル分解処理により形成される。

セル分解処理をマップ f でおこなうと、

$$f: B^n \rightarrow \cup_i \cup_{g_i} B^{u_i}$$

このマップは、既存の n 次元セル B^n を、セルの排他和にマッピングする。そのアタッチングマップ g_i は与えられている。そのアタッチングマップ g_i は全射であり、連続性を示す識別式

$$g_i: \partial \dots \partial B^n \rightarrow B^{u_i},$$

を満たす。ここで、 δ は $n - u_i$ 回繰り返される。

この状態モデルは、以下ようになる。

(a)セル分解: sセル B_s^p をネットワーク関連の S 社のサイト、cセル B_c^q を金融関係の c 社のサイト、mセル B_m^r をコンピュータ関連の m 社のサイトとおく。これらのセルは、ネットワークの技術分野に関する同値関係に基づいて、ウェブ上のサイトから識別されたものである。そのセル分解の処理は、

$$f_s: B_s^p \rightarrow B_s^{p-t} \cup_g B_s^t,$$

$$f_c: B_c^q \rightarrow B_c^{q-t} \cup_h B_c^t,$$

$$f_m: B_m^r \rightarrow B_m^{r-t} \cup_k B_m^t,$$

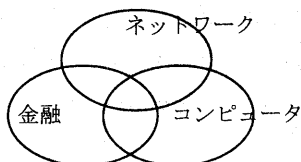
である。

ここで、g、h、k は、以下のような識別式である。

$$g: \partial \dots \partial B_s^p \rightarrow B_s^t (\theta \text{ は } p-t \text{ 回繰り返される}),$$

$$h: \partial \dots \partial B_c^q \rightarrow B_c^t (\theta \text{ は } q-t \text{ 回繰り返される}),$$

$$k: \partial \dots \partial B_m^r \rightarrow B_m^t (\theta \text{ は } r-t \text{ 回繰り返される})$$



(b)セル結合 : (a)で分解したセルを、以下のアタッチングマップにしたがって、結合する。

$p_m: \partial \dots \partial B_c^q \rightarrow B_m^t$ (∂ は $q \cdot t$ 回繰り返される),

$p_s: \partial \dots \partial B_c^q \rightarrow B_s^t$ (∂ は $q \cdot t$ 回繰り返される)

以上により、複数の技術分野にまたがって、同値関係によりセル分解、セル接合を達成することができ、その情報をデータベースに保存することが可能となる。

また、このデータは、同値関係を定義づけるアタッチングマップにより最適化することが可能である。

(ii)実装

本モデルを実装するために、特別なインターフェースを備えた XML 構造を準用する。前述したように、XML は XML ファイル中に XML データを内包するデータベースを形成する。しかし、その中から特定の情報を抽出する際には、XSLT のようなスタイルシートを用い、データの形式を再定義する必要がある。これは、その生い立ちが、文献情報管理からきていることによる。

この XML をセルのデータベースに活用するためには、XML データのスタイルを柔軟に定義できればよい。そのためには、ウェブから抽出された情報を、その検索情報についての XML タグを使用することにより、XML 形式のファイルに変換できるツールがあればよい。

今日、XQuery などの技術によって、インターネット上の多様な XML 形式のデータを効率的に選り出し、独自のデータ構造に変換する技術の構築がすすめられている。しかし、この手法のみでは、XML 形式で記述されていない情報をウェブから

抽出することは不可能である。

そこで、本論分では上記不具合を解決すべく、XML 形式で記述されていないデータの検索も可能にする手法を前述したセルモデルに従い、提案する。

(1) セル分解

- ① ウェブ情報の中から、求める“技術分野”の検索の対象となる部分を特定し、その“技術分野”のタグをつけて、XML データの階層構造(library.xml)を作成する。

<Computer>

Title : A Diffusion Model for Computer Animation of Diffuse Ink Painting

本論文のぼかしインク画のコンピュータアニメーションのための多次元的拡散モデルは実際の画像のもとと全く同じ分布の強さを作り得ることを証明した。共著者 : T.L.Kunii, G.V.Nosovskij, T.Hayashi

</Computer>

(2) セル結合

- ② ①の対象の中から、求める“キーワード”をもつ部分を特定し、その“キーワード”のタグをつけて上記ファイル上に追記する。

<Computer>

<title>Analyzing Human body Motions by Extracting Regions in the Configuration Paths : With an Application to Analyzing Skills of Shorinji Kempe</title>

<para>

本論文のぼかしインク画のコンピュータアニメーションのための多次元的拡散モデルは実際の画像

のものと全く同じ分布の強さを作り得ることを証明した。:</para>

```
<author>T.L.Kunii, G.V.Nosovskij,
T.Hayashi</author>
</Computer>
```

(3) 情報検索の例 (XQuery を使用)

③ ①、②で作成した XML データの階層構造 (library.xml)の中から、XQuery の照会の対象となる部分を特定し、後続の処理のため変数にバインドする。

著者が“TLKunii”である論文のタイトルを照会し、結果を result 要素に埋め込む場合、Xquery では次のように記述する。

```
<result>
{
For $b IN document("library.xml")//thesis
WHERE $b/author = "TLKunii"
RETURN $b/title
}
</result>
```

④ ③の結果を照会条件に従ってフィルタリングする。

⑤ フィルタリング結果の構造を変換しなおして、結果ツリーを作成する。

このような処理によって、求めるサイバースペースの構築が可能となる。

[参考文献]

[1] T. L. Kunii, "Creating a New World inside Computers -Methods and Implications-", Proc. of the Seventh Annual Conference of the Australian Society for Computers in Learning in Tertiary Education (ASCILITE 89), G. Bishop and J. Baker (eds.), pp. 28-51, Gold Coast, Australia, December 11-13, 1989; [also available as Technical Report 89-034, Dept. of Information Science, The University of Tokyo].

[2] T. L. Kunii, "The Architecture of Synthetic Worlds", Cyberworlds, T. L. Kunii and A. Luciani (eds.), pp19-30, Springer, Tokyo, 1998.

[3] T. L. Kunii, "Homotopy Modeling as World Modeling", *Proceedings of Computer Graphics International '99 (CGI99)*, (June 7-11, 1999, Canmore, Alberta, Canada) pp. 130-141, IEEE Computer Society Press, Los Alamitos, California, U. S. A.

[4] T. L. Kunii, "Valid Computational Shape Modeling: Design and Implementation", *International Journal of Shape Modeling*, World Scientific, December 1999.

[5] E. F. Codd, "A Relational Model for Large Shared Data Banks," *Communications of the ACM*, Vol. 13, No. 6, pp.377-387, June 1970.

[6] J. H. C. Whitehead, "Combinatorial Homotopy I", *Bulletin of American Mathematical Society*, vol. 55, pp. 213-245, 1949.

[7] J. H. C. Whitehead, "Algebraic Homotopy Theory", *Proceedings of International Congress of Mathematics, II*, Harvard University Press, pp. 354-357, 1950.

[8] 中山幹敏、奥井康弘 "XML 完全解説", 2001