

研究論文：論文

土木建造物の命名支援システムの開発

石原恵子*, 石原茂和*, 長町三生*, 竹林征三**

* 広島国際大学人間環境学部, ** 富士常葉大学環境防災学部, 風土工学研究所

DEVELOPMENT OF NAMING SUPPORT SYSTEM FOR CIVIL CONSTRUCTIONS

Keiko ISHIHARA*, Shigekazu ISHIHARA*, Mitsuo NAGAMACHI* and Seizou TAKEBAYASHI**

*Hiroshima International University, 555-36, Gakuendai, Kurose-cho, Kamo-gun, Hiroshima, 724-0695, Japan

** Fuji Tokoha University, 325, Obuchi, Fuji, Shizuoka, 417-0801, Japan

Abstract: The civil construction in a particular community should be given a name that appropriately reflects the climate and culture of the area (expressed as "fudo" in Japanese) where they are constructed, which will add social and cultural value to it. From the view of Fudo Engineering, the namer may well be required to have all the knowledge of the fudo of the area, however, it is not easy to acquire the adequate knowledge. Here, we will propose to use our newly developed computerized consultation system, which uses an electronic encyclopedia, to get the fudo information effectively. When the user inputs a name of the place, this system searches all the descriptions about it from the encyclopedia. The descriptive text includes the nature environment, origin, history and so on. Then, it automatically extracts the nouns and the adjectives from the text, and makes a word list semantically related to the each extracted word. The word list helps the user choose a word or coin a new word conceptually related to the community as a name for the construction.

Keywords: Naming support system, Morphological analysis, Concept hierarchy

1. はじめに

「文明」と「文化」という言葉はよく対比して用いられる。「civilization」の訳語である「文明」は、高度な技術、そしてそれによって得られる物質的な所産を指す。一方、「culture」の訳語である「文化」は、目に見えない、精神的あるいは知恵の所産を指している。この分類で言えば、物理的存在である土木建造物は文明ということになる。しかし、言うまでもなく文明と文化は不可分なものである。土木建造物は社会や価値といった文化の中で必要に応じて作られるものであるから、この点でも文明に属する土木建造物と文化は不可分なものである。

一方、命名とはソシュールが言うところの「シーニュ＝表記するための記号そのもの」、「シニフィアン＝言語の媒介する表現」および「シニフィエ＝内在する意味」が一体となったような行為である。言語が文化の中で大きな重要性をもつことを考えると、具体物に対する命名という行為は、単なる記号を産出することではなく、作られる場所の文化、あるいは風土を投影するような名前を産み出すことでなければならないであろう。

1.1. 風土工学と命名

風土工学で整理された、土木施設存在についての3超越、すなわち社会的・時間的・空間的構造 [注1, 2] は、文明を構成する人工物一般に広くあてはまる。人工物は文化の中で使用されることによって意味を持つのであり、文化を構成するもの、すなわち思考、感情、衣食住、環境、制度などといっ

た物事はそれぞれ社会・時間・空間の要素を持つからである。風土工学の考え方を命名という活動にあてはめると、この3つの構造を文化としてとらえる、つまり、その命名対象が存在する場所の文化を知る必要がある。しかし、文明を含む文化についての記述を得るために、書物や対象地域の住人、あるいは関係者から文化を構成するものすべてを調査することは不可能である。

1.2. 百科事典からの名称シード抽出

我々は命名のプロセスを機械で支援することを考え、風土を含む文化を知る最初の手がかりとして最も効率のよい情報源として百科事典を選んだ。その理由は、百科事典には風土を含む文化が記述されていること、そのテキストが場所単位で構造化されていて検索できること、さらにテキストが電子化されていることの3つである。多くの百科事典には行政地域の単位ごとに地名のインデックスをつけて整理され、それぞれに故事来歴、土木事業、地場産業、自然環境、風物、などといった文化や文明のデータが記述されている。風土工学を踏まえた命名の第1歩としては好適である。

しかし、百科事典を引いただけでは、必ずしも名称としての確かな語が得られるとは言えない。なぜなら百科事典の記述は、文化や文明のさまざまな側面を表すためにキーワード的な単語が多い、キーワードの数が多くても名称として使うにはあまりに表層的であったり使い古されていたりする単語も多いからである。もう一つの理由は、百科事典には多くの面にわたって記述されているために、その中に含まれるキー

ワード的な単語、すなわち名詞や形容詞の数が非常に多く、これらを漏らさず取り上げることは人手によれば困難な作業となることである。

そこで我々は、形態素解析と概念（意味）階層に基づく整理という2種類の日本語解析処理を備えた命名支援システムを考案した。

1.3. 形態素解析

日本語研究の中で書き言葉や話し言葉の語彙調査や文字調査が行われている。国立国語研究所では昭和41年以来、新聞の語彙調査、漱石・鴎外の用語調査、高校教科書の語彙調査などが行われている[注3]。語彙調査では、形態素解析によって文を単位分割し、それをもとに品詞認定、語種認定などが行われる。

形態素とはそれ以上意味的に分割できない最小の記号のことである[注4]。例えば「川」という名詞は形態素である。一方、動詞である「話す」は/hanas/までが、また形容詞である「大きい」は/ooki/までが活用形に共通である語幹であり、それぞれ一つの形態素である。/a/, /i/, /u/などの動詞の活用語尾や/ku/, /i/, /kere/などの形容詞の活用語尾もまた、それぞれで一つの形態素となる。副詞と助詞は活用しないので、一つの語がそのまま形態素である。助動詞は活用し、動詞または形容詞に後続する。このように、文を形態素の単位に分割し、前後の接続関係や辞書を用いて品詞を認定することが可能になる。

本研究では百科事典の記述から文化や風土を表すキーワードの候補として名詞と形容詞を取り上げるので、形態素解析から品詞認定の一部までを行うことになる。多数の名詞および形容詞を抽出するための、機械による自動抽出プログラムについては3節で述べる。

1.4. 概念階層と発想支援

形態素解析によって、百科事典の記述から多数の名詞と形容詞が抽出されてくる。しかし上に述べたように、これらは風土との関連をもつとはいえ、大部分は名称としてそのまま使えるわけではない。ふさわしい名称を導くには、得られたキーワードから出発して、命名者が発想をふくらませる必要がある。

ここで我々が発想支援として考えたのは、意味による言葉の整理と発見という方法である。何か疑問のあることを辞書や事典の類で調べたとき、そこからさらに調べていくと思いがけず面白い発見をする経験は多い。しかし、風土に関する言葉から事典辞書を引いていけば、名称としてよい言葉を発見できるとは限らない。辞書が電子化され高速に検索できるようになったとはいえ、やみくもに引き続けるのでは情報の海に遭難することにもなる。

「編集」というキーワードで思考のプロセスを考察した松岡は、編集の中に収集・分類などデータ操作に関わる「編集(compile)」と意味情報にさまざまな加工を施す「編集(edit)」という2つのプロセスがあるとしている[注5]。我々の発想支援では意味もしくは概念において言葉を「分類」する。一般の国語辞典では言葉の意味が詳細に記述されているが、

その言葉がどんな言葉のグループに属するかという分類作業には使えない。そこで我々は、言葉の意味の階層関係を記述した辞書を使うことにより、検討している言葉にどのような意味があり、どのグループに属するのかを調べることにした。そして、意味構造の中でその言葉が属するグループを探索し、同じグループに属する言葉には何があるのか、上位概念は何か、下位概念はどのように細かく分かれているのかを命名者に示して「発見」と「編集」を支援することにした。

我々の用いた概念階層辞書の中で、同じ意味階層にある言葉や、その上位概念あるいは下位概念を表す言葉には、地域性や風土性をより濃く表す言葉のほか、雅語も多く登場する。それらを使って風土工学で言う発想転換プロセスのうちの変化、拡大、たとえや代用に沿った言葉が発見できる。特に、たとえや代用のプロセスにある用・強・美、自然風土、歴史事象、概念にたとえる言葉は、その土地のよさを表し期待を込めた名称の候補になりうる。例えば「東広島市」についての百科事典記述から取り出した「フジ」は、単に一つの植物を表す名詞であるが、そこから概念階層を検索することで「藤波」という美しい言葉を見つけることができた[注6]。

1.5. コンピュータを用いた発想支援

本稿では、百科事典から命名のシードを抽出し、それに意味構造に基づく整理と探索を行うことにより、新たな命名の発想を支援するシステムについて述べる。風土工学に基づく命名のために、電子化された百科事典と、言葉を意味的な関係に基づいて整理した辞書である概念辞書を利用する。百科事典を地名で検索したテキストから単語を抽出すれば命名の最初の手がかりとして効率的である。しかし、そのまま名称とするにはあまりに表層的である場合も多い。そこで、抽出した名詞および形容詞から意味的なレベルに掘り下げて検討するための材料として概念辞書を利用する。ここで提案するシステムの目的は、このような大量の参考資料をひもとく作業の部分をコンピュータ・プログラムで自動化し、検索結果をユーザーに分かりやすい形で呈示し、ユーザーが考えを進めるごとに修正や追加が容易に行える環境を提供することである。このシステムの実現のために使われているテクニックは、計量国語学での日本語文法の統計的ルール、人工知能での自然言語処理、そして、風土および社会的存在からの命名技法であり感性工学の一分野を担う風土工学である。

2. システム全体の構成と命名プロセスの概要

この命名支援システムは2つのシステム、すなわち電子テキストからの命名シード抽出システム(図1)と、概念階層検索による発想支援システム(図2)から構成される。前者は百科事典CD-ROMの検索から得たテキストを自動的に名詞と形容詞を取り出してテキストファイルとして出力するので、後者は語のリストから概念階層辞書を検索して、その上位概念および同じ上位概念をもつ同列の概念を提示し、編集を促すものである。この2つを使うと、前者のシステムに地名を入力することでその土地についての百科事典記述から抜き出した名詞と形容詞を命名シードとして、これを後者の

概念階層検索による発想支援システムに入力して名称候補を引き出すこともできる。また、風土工学調査などを行って適切な命名シードが得られた場合は、これを後者のシステムに直接入力して発想を転換するために用いることもできる。

電子テキストからの命名シード抽出システムと、概念階層検索による発想支援システムの2つを連携させて用いる場合、命名プロセスは以下ようになる。

- (1) 電子テキストからの命名シード抽出システムの百科事典検索サブシステムで、命名対象の存在する地名を百科事典CD-ROMから検索する。これにより該当する土地に関する歴史・産業・文化などの情報が電子テキストの形で得られる。ここで得られた記述には、風土工学における命名技法の第1ステップ〔注2〕で挙げる名称素案、すなわち命名対象の風土を構成する空間軸風土資産、時

間軸風土資産および社会的風土における存在（土木施設の機能や姿、地域の期待）のうち、空間軸資産・時間軸資産に対応するデータが、名詞や形容詞の形で含まれている。そこで得られたテキストから形態素解析サブシステムを用いて名詞と形容詞を自動的に抽出する。

- (2) (1)で得られた言葉を命名シードとして命名に結びつけるためには、上で抽出された名詞・形容詞群の中から空間軸・時間軸資産に属する単語を選び出し、その意味から連想したり、発想を転換したりして、関連するさらにふさわしい言葉を探す過程が必要になる。出力された名詞・形容詞リストからユーザーが重要だと思うものを選んで概念階層検索による発想支援システムに入力すると、このシステムは意味的な分類の辞書を検索して、意味や上位・下位概念を提示する。また、ユーザーが意味的に同列にある語を検索して、命名シードをさらにふやすこともできる。

このシステムの主要部分はHyperCard (Apple Computer社)とOracle Media Objects (Oracle社)で開発した。システム全体はパーソナル・コンピュータMacintosh (Apple Computer社)のMac OS 9またはMac OS XのClassic環境で稼働する。

3. 電子テキストからの命名シード抽出システム

このシステムは、図1に示すように、百科事典検索サブシステムと形態素解析サブシステムから構成される。

3.1. 処理手順

以下の5ステップによって行う。

- (1) システムを起動すると、百科事典検索サブシステムが起動し、百科事典のCD-ROMをマウントするように促す。
- (2) 次に電子ブック・リーダーのプログラムが起動し、ユーザーが検索したい地名を市町村名で入力すると、百科事典から検索された記述が電子ブック・リーダーの検索結

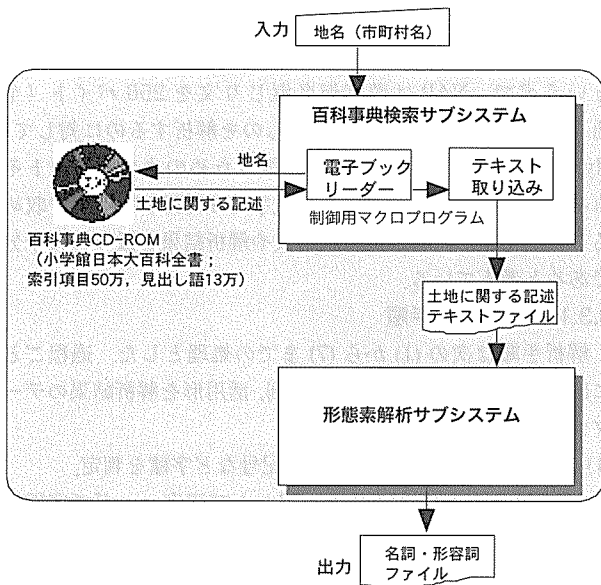


図1：電子テキストからの命名シード抽出システム

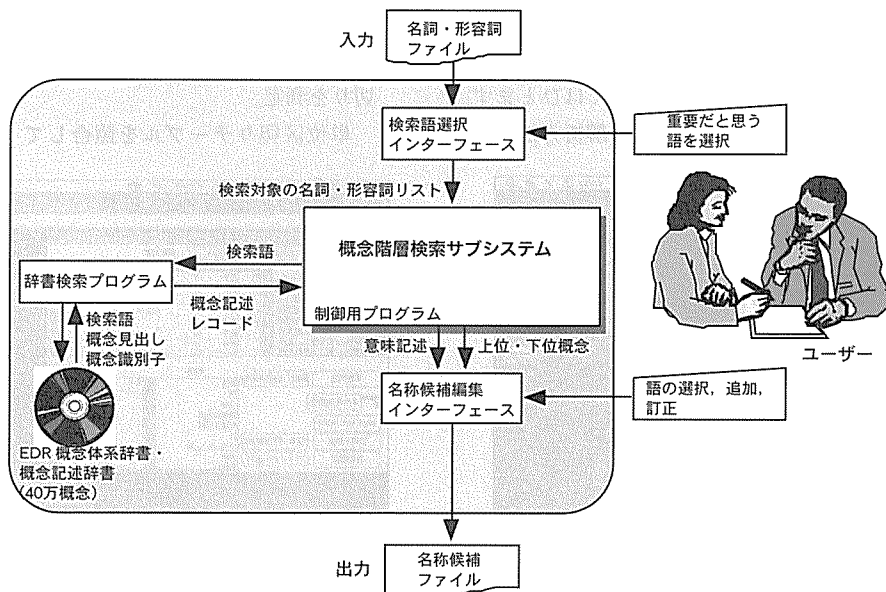


図2：概念階層検索による発想支援システム

びA(接続詞)/、「おける」を「/おけER(動詞命令形)/るPI(助動詞連体形)/」のように、区切り記号と品詞コード、活用コードを入れる。

- (4) 字種による統計的接続率・区切り率からの区切り判定。
対象とする文字の字種とその前後の字種を文字連続による単語分割の表と照合して、区切る、または区切らないを示す記号を入れる。
- (5) 活用形とのマッチングから品詞を認定。
語末の文字と字種についての語形ルールを適用し、品詞コードと活用コードを入れる。
- (6) 助詞・助動詞とその前にくる語のテーブルから区切りと品詞認定の訂正。

名詞や動詞と助詞・助動詞の接続規則について、助詞・助動詞接続テーブルと照合して品詞と活用形を訂正する。このテーブルには対象となる語とその品詞、その語の直前に用いることのできる助詞・助動詞、その語の直前に用いることのできる品詞と活用形、その2種類に一致しない場合に強制的に適用する品詞と活用形が記述されている。

- (7) 名詞と形容詞を抽出。

解析結果データシートから品詞が名詞あるいは形容詞(形容詞または動詞と判定されたものを含む)と判定されたものを、前の語との区切りまでたどり、それぞれ出力用フィールドに書き出す。さらにこれらをテキスト・ファイルに出力する。

3.3.2. 形態素解析の例

入力文はいずれもここで用いた百科事典、日本大百科全書電子ブック版(小学館)で地名を検索して得られた記述とした。

3.3.2.1. 解析例1

入力文：

新治(村) ニハル

群馬県北部、利根(トネ)郡にある村。三国(ミタ)山脈の南側、利根川の支流赤谷(アカヤ)川の流域を占め、国道17号(三国街道)が通じる。総面積の5%が耕地、78%が山林で、耕地では桑園がもつとも多く、古来養蚕が盛ん。第二次世界大戦後コンニャクイモ、葉タバコ、シイタケなどの生産が増えた。中心集落の布施(フセ)は旧宿場、下新田(シモシンデン)は塩原太助(シオバウラタケ)の出身地、猿(サル)ヶ京(キョウ)、湯宿(ユヅク)、法師(ホウシ)などの三国温泉郷で、三国街道のにぎわいを再現している。人口8153。〈村木定雄〉

抽出された名詞(全名詞44語のうち41語を正しく抽出、正解率93.2%)：

新治, 村, ニハル, 群馬県北部, 利根(トネ)郡, 村, 三国(ミタ)山脈, 南側, 利根川, 支流赤谷(アカヤ)川, 流域, 国道, 17号, 三国街道, 総面積, 5, %, 耕地, 78, %, 山林, 耕地, 桑園, 古来養蚕, 盛, 葉タバコ, シイタケ, 生産, 中心集落, 布施(フセ), 旧宿場, 下新田(シモシンデン), 塩原太助(シオバウラタケ), 出身地, 猿(サル)ヶ京(キョウ), 湯宿(ユヅク), 三国温泉郷, 三国街道, ぎわい, 再現, 人口, 8153, 村木定雄

抽出された形容詞(全形容詞2語のうち1語を正しく抽出、正解率50%)：

多く

抽出されなかった名詞(または誤って抽出された名詞)：

第二次世界大戦後コンニャクイモ, 法師(ホウシ), にぎわい
抽出誤りについて、それぞれ原因は次のように考えられる。

- (1) 漢字/カタカナおよびカタカナ/記号の接続は「区切らない」というルールがマッチした。
- (2) 「法師(ホウシ)などの」に対して「どの」が先にマッチしたため「法師(ホウシ)/な/どの」と区切りが入り、「な(助詞または助動詞)」の前は名詞と認定されなかった。
- (3) 「に(助詞)/ぎわい(名詞)/を(助詞)」という区切りがマッチしたため。

3.3.2.2. 解析例2

入力文：

東広島(市) ヒガシヒロシマ

広島県中部、西条(サイジヨウ)盆地の中心をなす市。1974年(昭和49)賀茂(カモ)郡西条, 8本松(ハチホンマツ), 志和(シワ), 高屋(タカヤ)の四町が合併して成立。山陽本線、国道2号、375号が通じる。古代には安芸(アキ)国府が置かれたとする説もあり、西条には安芸国分寺跡(国の史跡)がある。江戸時代には山陽道の宿駅四日市宿が置かれた。明治以降、賀茂郡の中心として県の出先機関などが置かれ、現在、県立農業短大、エリザベト音楽大学の分校などがあり、広島大学の統合移転計画が進行中である。また広島中央テクノポリスの中心として、電子工業やエレクトロニクスなど先端産業の誘致に努めている。現在、シャープや日本電気などの工場が立地している。このほか佐竹製作所の農機具は国内をはじめ東南アジアにも輸出されている。在来工業としては酒造業が有名で、西条の賀茂鶴(カモヅル)など多くの銘酒があり、白壁造りの酒蔵がみられる。松林が多く、マツタケの産地としても有名、三永(ミナガ)水源池はフジの名所として名高い。高屋地区の木原家住宅は江戸初期の豪商住宅で国の重要文化財に指定されている。人口8万4717。〈北川建次〉【本】『西条町誌』(1971・西条町)

抽出した名詞(全名詞85語中84語を正しく抽出、正解率98.8%)：

東広島, 市, ヒガシヒロシマ, 広島県中部, 西条(サイジヨウ)盆地, 中心, 市, 1974年, 昭和, 49, 賀茂(カモ)郡西条, 8本松(ハチホンマツ), 志和(シワ), 高屋(タカヤ), 四町, 合併, 成立, 山陽本線, 国道, 2号, 375号, 古代, 安芸(アキ)国府, 置, た, 説, 西条, 安芸国分寺跡, 国, 史跡, 江戸時代, 山陽道, 宿駅四日市宿, 置, 明治以降, 賀茂郡, 中心, 県, 出先機関, 置, 現在, 県立農業短大, エリザベト音楽大学, 分校, 広島大学, 統合移転計画, 進行中, 広島中央テクノポリス, 中心, 電子工業, エレクトロニクス, 先端産業, 誘致, 現在, シャープ, 日本電気, 工場, 立地, ほか, 佐竹製作所, 農機具, 国内, を, 東南アジア, 在来工業, て, 酒造業, 有名, 西条, 多く, 銘酒, 白壁造り, 酒蔵, 松林, マツタケ, 産地, 有名, 三永(ミナガ)水源池, フジ, 名所, 高屋地区, 木原家住宅, 江戸初期, 豪商住宅, 国, 重要文化財, 人口, 8万, 4717, 北川建次, 本, 西条町誌, (1971, 西条町)

抽出した形容詞(全形容詞4語中2語を正しく抽出、正解率50%)：

多く, 名高い

処理時間については、ここで解析した3つの例、すなわち新治村(271字)、東広島市(536字)それぞれ20秒および49秒であった(Apple Computer社Power Macintosh 9600, CPUはPowerPC 604e 235MHzを使用した)。使用する条件によっても異なるが、十分実用に耐える速さであると考えている[注10]。

4. 概念階層検索による発想支援システム

百科事典の記述から取り出した名詞および形容詞は、その土地に関連する地名、歴史、産業、風物などを含んだものである。これらを直接、土木施設の命名に用いるだけでなく、その土地のもつ風土や期待などに変換・昇華するためには、命名する人（ユーザー）の連想や発想転換が加えられる必要がある。

そこでユーザーの連想や発想転換がさらに豊かに行われるように支援するため、このサブシステムは、自動処理とその結果をユーザーに分かりやすい形で提示して、ユーザーの選択と編集を促すことに主眼をおいている。

このシステムは図2に示したように、語の選択や編集のためのユーザー・インターフェースを持ち、概念階層辞書検索プログラムを制御する。

4.1. 概念の分類

我々の使う言葉が示す概念を意味的に分類すると、包含関係あるいは階層関係として分類することができる。階層的な関係とは上位・下位の関係であり、上位の概念とはその概念を含んだ、より抽象的な概念である。下位の概念とはその概念のさらに具体的な概念である。例えば、「学校」という概念の上位概念には「組織」、「建物」、「機能」などがあり、下位概念としては「小学校」や「大学」などが考えられる。1つの概念が2つ以上の上位概念や下位概念を持つ場合もあるため、完全な木構造にはならないが、上位・下位の概念を順にたどっていくことで、意味的に抽象化されたもの、あるいは逆に具体化されたものを見つけ出すことができる。さらに、1段階上位の概念から見て同列に位置する概念を拾い集めると、その中に類義語や関連語を発見することができる。類義語辞書やシソーラスといった辞書が、会話や作文の際に、適当な、あるいはより適当な表現を求めたり、同じ表現を繰り返すことによる単調さを避けるために使われたりするように、概念階層の中で同列の概念は、ある種の発想転換を実現することができる。

この考えに基づいて、我々の発想支援システムでは、概念階層を検索して上位概念を表す語および同列に位置する概念を表す語を呈示することにより、ユーザーの発想転換を支援する。

4.2. システムの実装方法

4.2.1. EDR 概念体系辞書と概念記述辞書

このシステムで利用した EDR 概念辞書を含む 6 種類の大規模な辞書から構成される EDR 電子化辞書は、基盤技術研究促進センターとコンピュータ・メーカー 8 社との共同出資のもとに、1986 年から 9 年間のプロジェクトによって作られたものである（販売元：日本電子化辞書研究所）。このうち、概念辞書は単語辞書に語義として導入された 40 万個の概念についての知識が記述されている [注 11]。

本システムでは EDR 概念辞書 ver. 1.5 (1995) のうち、概念見出し辞書（ファイル容量 41.8MB）と概念体系辞書（ファイル容量 15.2MB）の 2 つを利用する。それぞれの辞書に含まれるレコードの例を図 5 に示す。

4.2.2. 大規模データの検索手法

ある概念の説明を言葉から検索するためには、概念見出し辞書の中からその言葉を概念見出しにもつレコードを取り出す処理を行う。また、ある概念の上位概念となる言葉を取り出すには、まず、それを概念見出しとして持つレコードを概念見出し辞書から検索してその概念識別子を取り出し、次に概念体系辞書の中で検索して該当する上位概念識別子を取り出し、それを用いてさらに概念見出し辞書を検索して概念見出しを求める、という処理を行うことになる。つまり、1つの語についての上位概念を 1 つ取り出すためには、この 2 種類の大規模な辞書を延べ 3 回処理することが必要である。下位概念を求める場合も同じである。

このシステム本体は HyperCard 2.3-J (Apple 社) のスタックとして開発した。上述のように大規模データベースを何度も検索する必要があるため、この処理を高速化するために 2 種類の処理を HyperCard から呼び出して使用している。1 つは BBfindXcmd (柴田淳氏 作) である。これは HyperCard の外部関数で、大量のテキストデータから検索対象の言葉を含む行を非常に高速に抽出する機能をもつ。2 つめは AWK で、これは不定長データの扱いに優れた汎用のスクリプト言語である。ここでは jgawk (2.15.2+1.0+1.2.5) を用いてプログラムを作成した。この 2 つを用いて、まず BBfindXcmd により辞書からテキストを行単位で取り込み、jgawk のプログラムにより取り込んだテキストから必要な項

概念見出し辞書 CPH.DIC (41.8MB) の例：

レコード番号	概念識別子	英語概念見出し	日本語概念見出し	英語概念説明	日本語概念説明	管理情報
CPH02040850f3c75	""	材料 [ザイリョウ]	"materials for experiments"	実験用に供される物	DATE="95/6/6"	
CPH02040860f3c77	""	材料 [ザイリョウ]	"thematic material that is used to create artistic compositions"	芸術作品をつくり出す題材	DATE="95/6/6"	
CPH02040870f3c78	""	材料 [ザイリョウ]	"intelligence that affects the prices in market (especially the stock market)"	相場の動きに影響を及ぼす情報	DATE="95/6/6"	
CPH02040880f3c79	""	材料学 [ザイリョウガク]	"a study concerned with industrial materials, called material science"	工業材料についての学問	DATE="95/6/6"	

概念体系辞書 CPC.DIC (15.2MB) の例：

レコード番号	上位概念識別子	下位概念識別子	管理情報
CPC0011971	30f74d	0ab2e5	DATE="95/6/7"
CPC0011972	443e5d	0ab2e6	DATE="95/6/7"
CPC0011973	443e4a	0ab2e7	DATE="95/6/7"

図 5：概念見出し辞書および概念体系辞書のレコード構造

目のみを取り出すという2段階処理とした。jgawk プログラムは独立したプログラムで、AppleScript によるマクロプログラムで制御している。

4.3. 処理手順

この概念階層検索による発想支援システムは、次の手順で処理を行う。

(1) 名詞・形容詞リストのテキストファイルを読み込んで、その中からユーザーが重要だと思うものを選びせる。

語を選んでリストに入れ、「EDR 単語検索」を実行すると、リストに含まれる語をすべて EDR 概念見出し辞書から検索する。その結果として、図7の最上段の枠に、検索語を含んだ文字列を見出し語にもつレコードが「概念の説明」として表示される。ユーザーはこれを読みながら、最初の単語リストに入れた意図に最も近い語を選び出す。

(2) 選んだ語に関連する語（上位概念および、その同じ上位概念をもつ同列の語）を検索してユーザーに提示する。EDR 概念体系辞書・概念記述辞書を利用する。

「検索」ボタンをクリックすると、レコード番号から概念体系辞書でこれらの階層関係にあるレコードを選び出し、それを表す単語を概念見出し辞書の中から検索して表示する（図7）。

(3) (1) で入力した語に検索結果を加えたものからさらにユーザーが重要だと思う語を選び、また、その時点で思いついた語を加えたり、編集させたりして名称候補リストとし、テキストファイルに出力する。

「候補出力」を実行すると最初に選んだ単語を含めて、検索結果すべてが名称候補のリストとして表示される。ユーザーがこれだと思う語を選んで「検討用に残すもの」リストに入れる。また、これまでの概念検索の結果を見ながらユーザーが思いついた、新しい語をこの中に加える、あるいは現在ある語を編集することもできる。こうして、最終的な名称候補を絞りこんだり、命名者の発想転換の材料にしたりすることができる（図8）。

5. 利用例

5.1. 百科事典から命名シードを抽出して発想支援システムにより名称候補を決定した場合（事例1）

群馬県新治村での事例を示す。まず電子テキストからの命名シード抽出システムに“新治村”を入力すると、百科事典から3.3.2.1.節の解析例に示した記述と名詞および形容詞が得られた。

次に、出力された名詞・形容詞リストのテキストファイル概念階層検索による発想支援システムに入力する。ここでは、百科事典から抽出した「猿ヶ京」の「猿」および「温泉」を選んで上位概念、さらに上位の概念、および同列に位置する概念を検索した。

この例では、上位概念として次のものが検索された：

サル、泉、建物、公共的な建造物。

さらに上位の概念として：

霊長類、水をたたえた地形、建造物

が検索された。

同列の概念としては次のものが検索された：

言わ猿、孤猿、こけ猿、小猿、子猿、シシオザル、巴猿、白猿、outpost、痩せ猿、山猿、タイワンザル（以下、サルの種類23種）、泉、黄泉、鉱泉、清泉、泉源、盗泉、走り井、噴泉、湧泉、冷泉、霊泉、裂罅泉、沢泉、Bramantesque（以下、建築様式34種）、牢屋、建物、新館、診療所、塾舎、城閣、女郎屋、館、塔、番小屋、家並み、上物、メモリアルホール、ガーデンハウス、パレデナシオン、宇宙ステーション、営造物、公用営造物、社会資本。

この事例では名称候補として「巴猿（峡谷に鳴く猿、峡谷の猿の鳴き声、旅愁）」、「清泉」などを得た。

5.2. 風土工学調査によって抽出した命名シードを発想支援システムに入力して名称候補を決定した場合（事例2）

大分県大山町にある大山ダム周辺施設の命名を提案した例を示す。大山町は福岡県と熊本県に接し、筑後川の上流となる大山川が流れる、人口4000人の町である。約40年前よりNPC運動と呼ばれる米麦から梅・栗の果樹農業への転換を成功させ、以来、農作物およびそれを原材料とするユニー

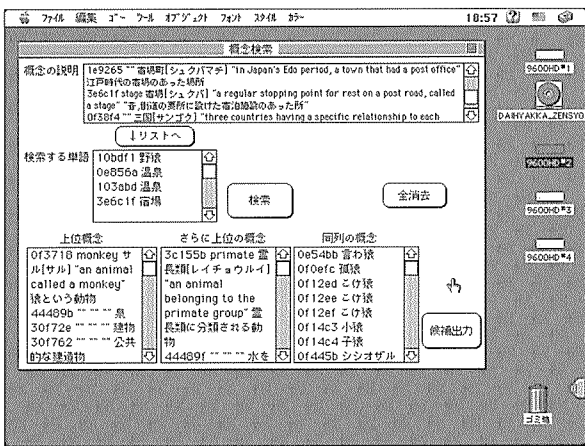


図6：概念検索結果を表示した画面

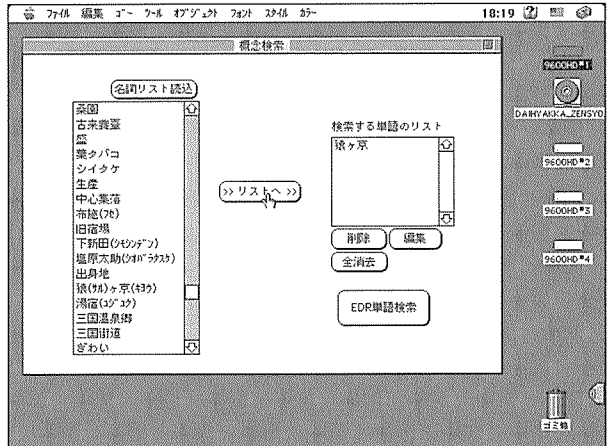


図7：検索した語から名称候補を検討させる画面

クな地元産品の企画および販売促進活動を今日まで展開していることで全国的に有名である。また、古くは弘仁年間(810)に記述された古文書などもあり、歴史伝承、自然風土ともに豊かである。風土工学に基づく調査を通して多くの命名シードとなる語が挙げられた。その中から「梅」および「竜」を取り上げて発想支援システムを利用し、名称候補を決定した。「梅」を検索語として同列の概念を検索すると、以下のものが得られた。

0e60f7 梅見[ウメミ] "to enjoy viewing ume blossoms" ウメの花を見て楽しむこと

0eb1c6 観梅[カンバイ] "the act of viewing plum blossoms" 梅の花を見て楽しむこと

1f479e 梅子[バイシ] "a fruit called an ume" 梅の実
この結果からダム周辺施設の名称として梅見橋、観梅橋、梅子トンネルなどが考えられる。

また、千丈淵にある竜神伝説から「竜」を検索語として検索した結果から、以下のものを見つけることができた。

1f3d26 伏竜[フクリュウ] "a person who has made remarkable achievements who is unknown to the world" 世間に知られていない優れた人物

1f3d27 伏竜鳳雛[フクリュウハウスウ] "a person who has made remarkable achievements and a hopeful youngster who is unknown to the world" 世間に知られていない優れた人物と有望な若者

後者に関連する語は次の意味を持つ。

1087f4 "" 鳳雛[ハウスウ] "a youngster who shows potential in some area" 前途有望な若者

この結果から、大山町の将来への希望を込めて鳳雛橋、鳳雛トンネルもふさわしいと考えられる。

6. おわりに

本稿では土木建造物に、建造する土地の風土や文化、将来への期待などを込めた名称を与えて、文化的な価値を高めることを考え、命名を支援するためのコンピュータ・システムを提案した。このシステムの前半処理である、名称シードを取り出すプロセスに、建造する土地の歴史や文化、産物など包括的に記述された電子百科事典の記述を用いることにより、その土地についての広い知識の中から命名につながるキーワードを拾い上げることを容易にすることができた。また、百科事典から抽出した語からでも、風土工学に基づく綿密な調査から抽出した語からでも、概念階層の中を検索することにより、連想したりイメージを膨らませたりする、ある種の発想転換を助けることができた。利用例に示した、地名の一部、あるいは動物としての「猿」から旅愁の意味をも持つ「巴猿」を、産物あるいは植物としての「梅」から「観梅」の楽しさ、すがすがしさのイメージを、また、東広島市を代表する花である植物の「フジ」からは一面に咲き誇る藤棚にわたる涼風を思い起こさせる「藤波」を得ることができた。これらの例は、提案したシステムを使うことにより、命名者にとっては、大量のデータを検索する作業の部分の労力を軽

減するとともに、効果的な思考材料を参照しながら、類義語辞書やシソーラスと同等もしくはそれ以上に豊かな意味を持つ名前を、導き出せる可能性を示すものであると考える。

謝辞

国立国語研究所・言語体系研究部長(当時)の故 中野洋先生には形態素解析およびその周辺プログラムのソースリストをご提供いただいた。BBfindXcmdの作成は柴田淳氏による。プログラムの使用について快諾をいただいた両氏のご協力に深謝いたします。CeDar 1.2.2の作成は川上三郎氏による。GNUawkの作成、日本語化、MacOSへの移植はそれぞれFree Software Foundation、田中良知氏、山下巖氏による。

注

- 1) 長町三生, 竹林征三: 土木施設の命名技法に関する風土・感性工学的研究, 土木計画学研究・講演集 19, 1, 301-304, 1996.
- 2) 竹林征三: 風土工学序説, 技報堂出版, 1998.
- 3) 中野洋: パソコンによる語の認定処理, 国立国語研究所報告 103, 研究報告集 12, 秀英出版(大日本図書), 83-130, 1991.
- 4) 渡辺実: 3章 品詞分類, 大野晋, 柴田武(編), 岩波講座日本語6文法I, 岩波書店, 85-128, 1976.
- 5) 松岡正剛: 知の編集工学, 朝日新聞社, 1996.
- 6) 石原茂和, 石原恵子, 長町三生, 西野達夫, 竹林征三: 土木建造物名称の決定支援システム, 第2回感性工学会大会予稿集, 112, 2000.
- 7) 電子ブック版日本大百科全書, 小学館, 1996.
- 8) 淵 武志, 米澤明憲: 日本語形態素解析システムのための形態素文法, 自然言語処理, 2, 4, 37-65, 1995.
- 9) 中野洋: パソコンによる日本語研究入門, 笠間書院, 1996.
- 10) 石原恵子, 石原茂和, 長町三生, 竹林征三, 西野達夫, 市坪誠, 小松孝二: 電子ブック百科事典からの名詞/形容詞自動抽出システム Anexee の開発, 広島中央女子短期大学紀要, 35, 1-9, 1998.
- 11) 日本電子化辞書研究所: EDR 電子化辞書, 概念辞書, 1995.