

## 仮想電子辞書の設計と実現

3 R - 3

芳野学 都司達夫 宝珍輝尚  
福井大学工学部

### 1 はじめに

電子辞書の存在意義はその文法構造に基づく説明項目の配置の強制を解放し、辞書が潜在的に保有している豊富で多角的な情報をユーザーの望む配列で提示することが出来ることであると考えられる。電子辞書を含めた電子出版物の構造化には、たとえばEPWING[1]などの規格が存在しており、それに従った電子辞書も出回りはじめています。また、既存辞書のハイパーテキスト化の方式なども提案されている[2]。しかし、辞書の論理構造(スキーマ)を前面に押し出しユーザーレベルで既存辞書からユーザーニーズにかなった辞書を定義し、利用するためのシステムは見当たらない。そこで、本稿では上記のような目的を達成するための「仮想電子辞書」の設計とその実現について述べる。

### 2 仮想辞書について

仮想辞書には二種類ある。仮想部分辞書は単一の辞書からユーザーが必要とする情報を、ユーザーが望む形で再配置して得られる単一辞書であり、仮想結合辞書は、一ないし複数の辞書から必要情報を「結合項目」を介して、結合して得られる単一辞書である。

#### ● 結合項目

例えば、英語シソーラス辞書において検索したい見出し語の類義語を検索する時、その日本語の意味も同時に知りたい場合、まず見出し語で英語シソーラス辞書中の英語の『類義語』を検索し、この『類義語』を見出し語として、英和辞書の『語義』を検索することになる。しかし、あらかじめこのような検索を目的として定義された仮想電子辞書では、これらの項目(『類義語』と『語義』)を結合することによって新たな仮想辞書として「和訳付き英語シソーラス辞書」を定義する(図1)。これらの結合される二つの項目(『類義語』と『語義』)を結合項目と呼ぶ。

#### ● 結合条件と結合可能性

辞書項目同士の結合を行うにあたって他の辞書項目の対応関係を結合条件として指定することで、冗長な情

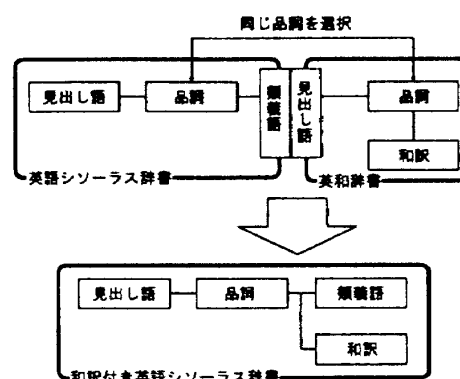


図1: 結合項目と結合条件

報を排除し、より目的にかなった辞書の定義を行う。例えば上記の例の場合、「英語シソーラス辞書の類義語と英和辞書の語義の品詞は同じ」であるという条件を課さなければ、異なる品詞同士が結合することになる。辞書項目のデータ形態には文字種(ひらがな、英字、漢字など)の違いや単語または単語列かにより、様々なものが存在するが、異なるデータ形態の辞書項目同士は結合することが出来ない。また、一部のデータ形態はそれを加工することによって他のデータ形態と結合することが可能である。

### 3 仮想辞書の定義と検索

例として図1の和訳付き英語シソーラス辞書の定義と検索例を図2に示す。DEFINE文により、仮想辞書の辞書名とその論理構造を定義し、FROM句により、定義のドメインとなる母体辞書群を指定する。次に仮想辞書の辞書項目が母体辞書のどの辞書項目に対応付けられるかをASSIGN文で指定する。辞書項目は、辞書名から目的の辞書項目までの項目に至るパスとして記述する。WITH ENTRIES句では仮想辞書の見出し語集合を指定する。指定された見出し語に対しては高速検索のための索引がシステムにより作成される。さらに、CONNECTED BY句では結合項目が指定され、必要に応じてwhere句により結合の条件が指定される。仮想辞書の検索はFIND文により行われる。検索条件をWHERE句で記述して、検索の絞り込みを行うことができる。

#### ● 仮想辞書定義と検索の例

例中 EJ は英和辞書、ET は英語シソーラス辞書である。

```

DEFINE 和訳付き英語シソーラス辞書 (見出し語,
見出し項目 (品詞, 類義語項目 (類義語, 和訳)))
FROM EJ, ET AS
ASSIGN
見出し語 = ET. 見出し語
見出し項目. 品詞 =
ET. 見出し項目. 品詞項目. 品詞
見出し項目. 類義語項目. 類義語 =
ET. 見出し項目. 品詞項目. 類義語項目. 類義語
見出し項目. 類義語項目. 和訳 =
EJ. 見出し項目. 品詞項目. 語義
WITH ENTRIES 見出し語, 見出し項目. 品詞
CONNECTED BY
ET. 見出し項目. 品詞項目. 類義語項目. 類義語,
EJ. 見出し項目. 品詞項目. 語義
WHERE
ET. 見出し項目. 品詞項目. 類義語項目. 類義語 =
EJ. 見出し項目. 品詞項目. 品詞
    
```

```

FIND 和訳付き英語シソーラス辞書
("test", 見出し項目 ("動詞", 類義語項目))
    
```

[検索結果]

```

見出し語: test
品詞: 動詞
類義語 (1): check
和訳 (1): 1. 阻止する, 防ぐ, 抑制する
           2. テストする, 検査する, 調べる
           3. 預ける 4. 王手をかける
和訳 (2): 1. 一致する, 符合する
           2. 検査する
           3. 小切手を振り出す
類義語 (2): example
和訳 (1): 1. 調べる, 検査する, 審査する
           2. 試験する
           3. 診察する, 検診する
           4. 尋問する
    
```

図 2: 和訳付き英語シソーラス辞書の定義と検索

● 非構造検索

通常、辞書の検索は見出し語から始まる辞書構造に沿って行われるのが普通である。しかし目的の辞書に対して、ある条件の下に全文を縦断的に検索をしたいという要望も存在する。本システムではこのような検索を非構造検索と呼び、辞書項目同士を縦列結合することにより、このような見出しによらない検索を目的とした仮想辞書の定義およびその検索を可能にしている。

● 仮想辞書の実体化

仮想辞書は、母体辞書に対する「ビュー」であるという性格上、検索コストのオーバーヘッド、特に、検索時の結合処理コストが高い。そこで、本システムでは、汎用性が高く、頻繁に使用する可能性のある仮想辞書

に対して実体化処理を行うことでこのようなオーバーヘッドを回避する。

4 仮想辞書システム

仮想辞書システムのクライアントは (1) ユーザからの仮想辞書定義や、仮想辞書検索の指示を受け、それぞれを DEFINE 文、FIND 文に変換する GUI、(2) DEFINE 文、FIND 文を構文解析し、対応する内部データ構造を 1 次記憶中に作成するパーザ、(3) 内部データ構造を参照しながら仮想辞書操作を行うインタープリタ、などからなる。また、サーバは (1) 母体辞書 DB、母体辞書定義文 DB、仮想辞書定義文 DB、(2) クライアントからの要求にしたがって、実際に母体辞書の検索を行うライブラリ、(3) 仮想辞書実体化サブシステム、などからなる。

仮想辞書定義 GUI では、ユーザは仮想辞書の母体辞書を選択し、その辞書構造の中から必要な項目を選択し、それを仮想辞書の辞書項目に対応付けながら仮想辞書の定義を行う。作成された仮想辞書定義文 (DEFINE 文) はサーバ側の仮想辞書定義文 DB に登録される。GUI において仮想辞書がユーザにより選択されると、その時点でパーザはサーバ側の仮想辞書定義文をロードし、仮想辞書定義データ構造を作成する。仮想辞書検索文は GUI で生成され、やはりパーザにより内部データ構造に変換された後、インタープリタが定義データ構造を参照しながら、サーバ側の母体辞書検索ライブラリルーチンを適宜呼出し、その検索結果をフォーマットしてユーザに提示する。

5 おわりに

仮想辞書システムの提案を行い、その概要について述べた。現在、本システムを実装中である。

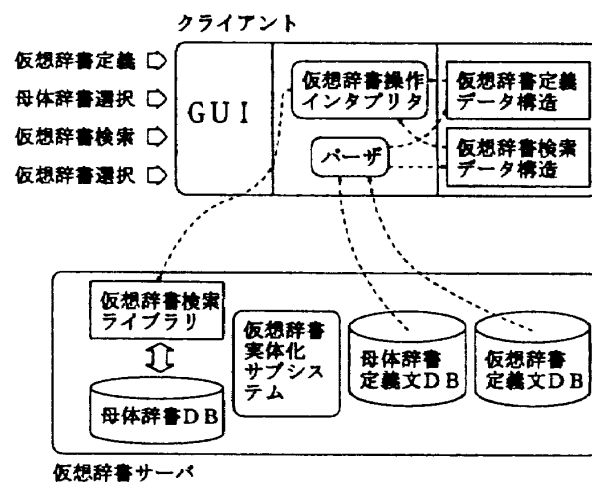


図 3: 仮想辞書システム

文献

- (1) EPWING コンソシアム: <http://www.epwing.or.jp>
- (2) 内藤、山下、松山、榎木: オンライン辞書のハイパーテキスト化手法、情処学論、34, 2, pp.320-329, 1993.