

意味フレームの階層化ネットワーク (HFN) を

いかに構築するか

〈 x が (z から) y を守る〉の HFN の構築の実例を通じて

黒田航

(独) 情報通信研究機構

けいはんな情報通信研究センター

1 はじめに

この論文では [3] などで紹介されている階層化ネットワーク (Hierarchical Frame Network: HFN) を構築するための手法を紹介する。簡単に概要を説明すると、カード分類を実践すればよいということである。これは誰にでも可能な手法である。

2 〈 x が (z から) y を守る〉の HFNA

私は以下に紹介する「手を使った」方法で〈 x が y を z から守る〉の階層化ネットワークを構築した。頭だけでなく、手を使うというのは実は重要な要素である。

2.1 基本用法の列挙

まず日英対訳コーパス¹⁾の日本語部をコーディングし、それを元に基本的用例の列挙を試み、次のような語群を得た:

- (1) 1. リードを守る; 2. (野球の) ショートを守る; 3. 景観を守る; 4. 権利を守る; 5. 秩序を守る; 6. 地位を守る; 7. 家を守る; 8. 教えを守る; 9. 基本を守る; 10. 自由を守る; 11. 国を守る; 12. 伝統を守る; 13. 自然を守る; 14. 権利を守る; 15. リードを守る; 16. 野生動物を守る; 17. 法律を守る; 18. 憲法を守る; 19. 友人を守る; 20. 自分(の身)を守る; 21. リードを守る; 22. ナワバリを守る; 23. リードを守る; 24. 墓を守る; 25. 約束を守る; 26. 目を守る; 27. 規則を守る; 28. 巣を守る; 29. 子供を守る; 30. 生活を守る; 31. 国境を守る; 32. 家族を守る;

この語群は完全にコーパス解析によって得られたものではない。一部は作例によって補われている。

2.2 幾つかの注意

2.2.1 注意 1

重要なのは、この基本的用例の列挙の段階で十分な用法の広がりを持つことである。用法の広がり全体像をつかむには、何といてもコーパスの実例に先に当たることをお勧めする。十分に訓練を積んでいないと、直観だけで用法の全体を取りだすのは至難の業であるし、仮にそれが可能だとしても、非常に非効率的な作業になる可能性がある。大量のデータを相手にする場合、効率を追求するのは、重要である。

2.2.2 注意 2

これらが本当に基本例であるという保証はない。それは多変量解析のような明示的操作が必要である。だが、始めからそれをする必要はないだろう。それは可能だが、言語学というより工学だろう。これが言語学であるためには、基本例を見抜き、それだけを取りだすための意味直観に基づいて行われる必要があるだろう。これがおそらく、もっとも研究者の資質が問われる部分であろう。

2.3 カードを使った用例の分類

(1) の語群には、ある程度の秩序が認められるが、完全なものではない(実際、(1) の語群十分に乱雑化させなかった)。

2.3.1 トークンとしての分類カードの作成

このような表現群を POST IT! に一つずつ書く。これをカード群 C とする。

POST IT! の代わりに普通の紙のカードに書いてもよいが、POST IT! だと壁や机に貼られて剥がせて便利である。

先ほども強調したことだが、用例をトークン化 = 記号化し、それらのトークンの集合の要素に内在す

¹⁾ <http://www2.nict.go.jp/jt/a132/members/mutiyama/jea/>

る類似性という抽象的な距離を、実際の空間的距離に対応させるという具体化の操作は—個人差がないわけではないものの—多くの人々が特別な訓練を必要とせずに驚くほど簡単に行なえる問題解決の手法であり²⁾、HFNA構築の強力な助け船になるはずである。

2.3.2 グループ化の手順と秘訣

カードのセットCが準備できたら、いよいよCの要素を類似性に基づいて、小さいグループに分けることになるが、これには幾つか重要な秘訣が存在する。例えば、

- (2) a. **ボトムアップ方式**でない、絶対に分類はうまくゆかない。**トップダウン方式**の分割がうまく行くのは、非常に規模の小さい、些細な場合だけである。従って、Cをまず C_1 と C_2 の二つに、次にいずれか一方を更に二つにという風に分割するのは、**絶対にうまくゆかない**。
- b. その代わり、なるべく意味の近いもの同士を小さなグループにしているうちに「島」ができ、そういう「島々」が結合して、段々と大きくなってゆくうちに自然にネットワークが形成されるというやり方が推奨される³⁾。
- c. 幾つかの小さいグループが統合されることになるが、その際に小さいグループを保存するようにする。
- d. 小グループの結合に際しては、用例の小グループの区切りを保存するだけでなく、それらの間の「距離」—意味の近さ、遠さ—を保存するように心がける。
- e. この際、同一の表現(例えば「国を守る」)が分類のあちこちに現われる可能性を排除してはイケナイ。**分類の部分的重なりは自然な現象である**。決して異常なことではない。
- f. 分類しながら類似例を新しく追加するのは、しばしば効果的である。これが可能であるということは、正しい分類を行

なっている見こみを与える。

(2e)のは特に注意を促しておきたい。多くの研究者が分類をキレイにするため(だけ)に、無理やり分類をツリーにしようとする。だが、ツリーは**多重分類**の可能性を排除し、結局は錯誤に導く。**一つの表現が一つの分類的意味をもつとは限らない**⁴⁾。

以上の注意の下で、分類の秘訣を紹介しよう。

2.3.3 類似性臨界内要素結合法を用いたボトムアップなグループ化

私が便宜的に**類似性臨界内要素結合法**と呼ぶボトムアップなグループ化の手順はおおよそ以下の通り:

(4) 類似性臨界内要素結合法の定義:

- a. 抽象的な距離と用いた類似性の定義:
Cのどの二つの用例 c_i, c_j を取ってきても、それらはある一定距離 $d(c_i, c_j)$ だけ離れている。
- b. 類似性臨界の定義:
 c_i, c_j が同じグループであると見なすための条件は、それらのあいだの距離 $d(c_i, c_j)$ が、ある適当な値 d_k より小さいことだとする。
- c. 類似性臨界内性の定義:
 $d(c_i, c_j) < d_k$ の場合、 c_i, c_j は(お互いの)類似性臨界内にあると言う。(二つの用例 c_1, c_2 が同一である場合、類似性臨界距離は0である)。

類似性臨界距離を(0から)段々大きくしてゆくことで、再帰的にグループ化を得るのが類似性臨界内要素結合法である⁵⁾。

2.3.4 類似性臨界内要素結合法のイメージ化

参考までに類似性臨界内要素結合法のイメージを図1, 2, 3に示しておく。もちろん、臨界距離は数値化する必要はなく、直観的でよい。重要なのは、それを一定に保つということである。異なる類似性臨界距離の分類をごっちゃにするのは、人文系の研究者の代表例としての言語学者の分類の悪いクセである。

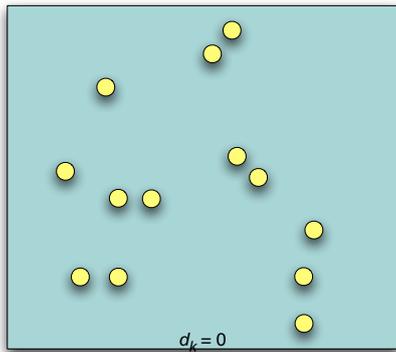
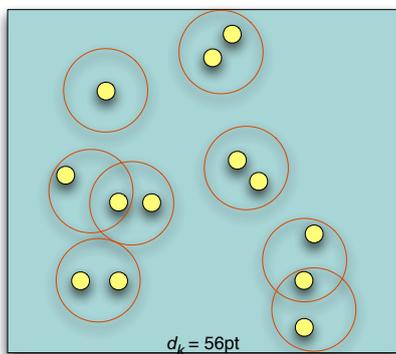
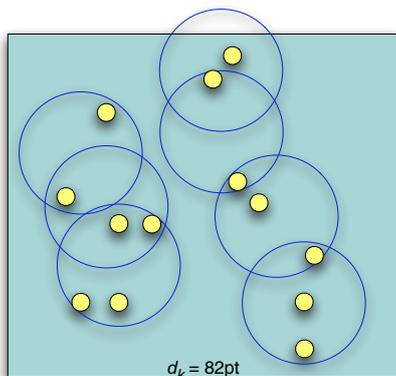
実際の操作をもう少し具体的に言えば、次のようになる:

²⁾ この能力を、私は生得的なものではないかと怪しむほど、この手法を使って多くの人々が訓練なしに高度な分類が可能である。この話は、古くはC. Lévi-Straussの『野生に思考』に「器用仕事」(bricolage)という名で登場する能力の起源の問題でもある。

³⁾ これは単なるテクニックではなく、ヒトの脳は実際にそうやって事例を分類、整理していると信じる理由が充分にある。

⁴⁾ 私たちが素性ラティスを想定する最大の根拠はここにある。意味素性をうまく使わないと、このような状況はうまく表現できないし、ツリー構造ではなくラティス構造だと考えないと、うまく表現できない。

⁵⁾ 多変量解析の知識のある方にはもうおわかりだと思うが、この手法は単にクラスター分類(clustering)の別の見方にすぎない。

図1 初期 ($d_k = 0$)図2 第一臨界距離によるグループ化 ($d_k = 56\text{pt}$)図3 第二臨界距離によるグループ化 ($d_k = 82\text{pt}$)

(5) a. 小さ目に設定した一定の類似性の距離の範囲にあると感じられる (つまり類似性の「臨界」内部にある) ものだけで微小なグループ $G = \{g_1, g_2, \dots\}$ を作る. ただし, 類似例のない, 「孤立した」用例 (e.g., “基本を守る”) は, ムリに何かのグループに入れなくてよい.

b. グループ G ができたら, 類似性の許容度をあげ, 臨界範囲を大きくする. この臨界の範囲内で近いと感じられるものだけで上位グループ $G' = \{g'_1, g'_2, \dots\}$ を作る.

これで, 例えば, $G: \{g_1, g_2, \dots\} \Rightarrow G': \{\{g_3, g_6\}, \{g_2, g_1, g_4\}, \dots\} (= \{g'_1, g'_2, \dots\})$ のようなグループ化を得る.

(6) (5) のサイクルを繰り返し, 上位グループを「孤立した」用例がなるなるまで大きくし, それが満足されたら, いったん止める.

概念ネットワークの複雑性にもよるが, おそらく, 二, 三回程度拡大すれば, (6) の終了条件は満足され, 目的に達する.

2.4 グループ化の結果: 第一段階

このような点に注意しながらグループを結合してゆけば, 例えば, 次のような四つの中規模グループが得られるだろう:

- (7) a. 〈生物が自分の体 (重要な部分, 特に攻撃に弱い箇所) を守る〉
- i. 自分 (の身) を守る;
 - ii. 目を守る; 頭を守る; ??心を守る
 - iii. 弱点を守る;
- b. 〈生物の個体が自分と利害を共有する他の重要な個体 (e.g., 〈家族〉, 〈友人〉) を守る〉
- i. 子供を守る; 家族を守る; 家を守る
[+metonymic];
 - ii. 友人を守る;
 - iii. 恋人を守る;
 - iv. 町を守る [+metonymic];
 - v. 国を守る [+metonymic];
- c. 〈生物個体が重要な場所/空間/範囲を守る〉
- i. 巣を守る; ナワバリを守る; 家を守る; 墓を守る; 町を守る; 国を守る;
 - ii. 国境を守る; 陣地を守る; 砦を守る; ???野原を守る;
 - iii. ショートを守る; マウンドを守る; *

- ピッチャーを守る; ゴールを守る;
- d. 〈主にヒトが重要な物事を守る〉
- i. リードを守る; 利益を守る; 地位を守る; 首位を守る; *ピリを守る; 家を守る; 生活を守る; 墓を守る;
 - ii. 基本を守る;
 - iii. 憲法を守る; 自由を守る; 権利を守る; 秩序を守る;
 - iv. 教えを守る; 掟を守る;
 - v. 法律を守る; 規則を守る; 憲法を守る;
 - vi. 約束を守る;
 - vii. 自然を守る; 野生動物を守る;
 - viii. 景観を守る; (平和) 憲法を守る; 伝統を守る;

ただし、これは第一段階である。理由は後で説明する。

2.4.1 第一段階から得られる知見

他の三つに比べ (11d) の内実は混乱しているが、この第一段階の分類結果からでも、およそ次のことが見て取れる:

- (8) 生物個体 x が y を守るのは、
- a. y が 〈 x 自身〉であるか、
 - b. y が 〈 x の〈家族〉、あるいは〈仲間〉のような特別な個体 x^* 〉であるか、
 - c. y が x か x^* にとって〈重要な場所〉、あるいは〈価値のある物事〉であるときである。
- (9) y は x にとって何らかの意味で〈重要なもの〉、〈価値のあるもの〉、しばしば〈利益をもたらすもの〉である。

以上の知見の下で、今度は (7) の第一段階のグループ化を精練させることを考える。

2.5 グループ化の結果: 第二段階

第二段階では次のことに注意しよう:

- (10) a. C のグループ化には、 C を幾つかのグループに「わける」プロセスと、個々のカードを「一緒にする」プロセスとの、二つの異なる側面がある。これらの二つの側面を区別するのは効果的である。
- b. 従って、最初の結びつきは中規模で、その後の中規模グループを大きなグループに統合すること。個々の中規模グループを、より詳しく、小グループに分割する

という二段構えの方針がよい。

- c. 個々のグループへの「ラベルづけ」を試みる。その際、意味素性を積極的に利用することが奨められる。

このような点に注意しながらグループを結合してゆけば、例えば、先ほどの四つの中規模グループは次のように細分化されるだろう:

- (11) a. 〈生物が自分の体 (重要な部分、特に攻撃に弱い箇所) を守る〉
- i. 〈自分の身の安全〉を守る; 自分を守る; 身を守る;
 - ii. 〈自分の弱点〉を守る: 目を守る; 頭を守る; ??心を守る;
- b. 〈生物の個体が自分と利害を共有する他の重要な個体を守る〉
- i. 〈家族〉を守る: 子供を守る; 妻を守る; 夫を守る; 孫を守る; 家を守る [+metonymic];
 - ii. (配偶者候補としての) 〈恋人〉を守る;
 - iii. 〈自分の属する共同体、あるいはそのメンバー〉を守る 〈仲間〉を守る: 友人を守る; 友だちを守る;
 - iv. 町を守る [+metonymic]; 国を守る [+metonymic];
- c. 〈生物個体が自分、あるいは自分と利害を共にする個体群にとって重要な場所/空間/範囲を守る〉
- i. 動物が守る
 - A. 〈巣〉を守る (cf. 〈家〉を守る);
 - B. 〈ナワバリ〉を守る;
 - ii. ヒトが守る
 - A. 家族規模: 〈家〉を守る (cf. 〈墓〉を守る);
 - B. 集団規模: 〈町〉を守る (cf. 〈古里〉を守る); 〈国〉を守る; 〈重要な場所の一部の特定の区画〉を守る: 国境を守る; ???野原を守る;
 - C. 〈重要な施設〉を守る 砦を守る; 陣地を守る;
 - D. 〈集団競技の特定のポジション〉を守る: ゴールを守る; ({バスケット; 野球; ???ラグビー}) でセンターを守る; ショートを守る; マウンドを守る; *ピッ

チャーを守る;

- d. 〈主にヒトが重要な物事を守る〉
- i. 主にヒトが〈資産 (非公利的)〉—特に個人の資産, 特定集団の資産—を守る: リードを守る; 利益を守る; 地位を守る; 首位を守る; *ビリを守る; 家を守る; 生活を守る; 墓を守る;
 - ii. ヒトが〈(共有) 資産 (公利的)〉を守る: 憲法を守る; 自由を守る; 権利を守る; 秩序を守る;
 - iii. ヒトが〈(共有) 資源 (公利的)〉を守る:
 - A. ヒトが〈自然資源〉を守る: 自然を守る; 野生動物を守る;
 - B. ヒトが〈文化資源/資産〉を守る: 景観を守る; (平和) 憲法を守る; 伝統を守る;
 - iv. ヒトが複数個体間の利益を保証するために〈取り決め〉を守る (遵守):
 - A. 基本を守る;
 - B. ヒトが〈取り決め (公利的)〉を守る: 法律を守る; 規則を守る; 憲法を守る;
 - C. ヒトが〈取り決め (小集団利的)〉を守る: 教えを守る; 掟を守る;
 - D. ヒトが〈取り決め (私利的)〉を守る (遵守): 約束を守る;

2.5.1 意味素性の有効利用

この段階になると, だいぶ HFNA らしくなって来る。だが, 中規模グループ = クラスター内部の階層性は明らかだが, それを一貫した形で HFNA にまとめ上げるのは, 直観のみでは難しい。何より怖いのは, 知らないうちに誤っている可能性がある。そのような問題を回避するため, [?private], [?public], [?temporal] のような**意味素性**をうまく用いて, **素性ラティス** (feature lattice) を構成することを考える。

2.5.2 幾つかの注意

細分化の際に, 概念メタファーによる写像 [1, 2] の効果は始めは無視した方がいい。重要なのは元領域の情報ではなく, 先領域の情報である。

3 HFN の構築

[未完]

3.1 グループ化を階層関係として表現する

HFNA の基本は, §2.3 の手順で得られたグループ化を階層関係として表現することである。

4 終わりに

この文書で私は, カード分類を実践し, 階層ネットワーク (Hierarchical Frame Network: HFN) を構築するための手法を紹介した。これにより HFN の構築が研究者にとって「黒魔術」ではなくなることを祈る。

参考文献

- [1] G. Lakoff and M. Johnson. *Metaphors We Live By*. University of Chicago Press, 1980. [邦訳: 『レトリックと人生』 (渡部昇一ほか 訳). 大修館.]
- [2] G. Lakoff and M. Johnson. *The Philosophy in the Flesh*. Basic Books, 1999.
- [3] 黒田航, 中本敬子, 金丸敏幸, 龍岡昌弘, 野澤元. 「意味フレーム」に基づく概念分析の射程: Berkeley FrameNet and Beyond. 日本認知言語学会第 5 回大会 Conference Handbook, p. 133, 2004.