

# 知識表現 KRNL による知識共有システムの構築

## Knowledge Sharing System based on A Knowledge Representation Scheme KRNL

友部 博教\*1      谷口 智哉\*2      松尾 豊\*3      石塚 満\*2  
 Hironori Tomobe      Tomoya Taniguchi      Yutaka Matsuo      Mitsuru Ishizuka

\*1名古屋大学情報科学研究科

Graduate School of Information Science, Nagoya University

\*2東京大学情報理工学系研究科

Graduate School of Information Science and Technology, The University of Tokyo

\*3独立行政法人産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

In this paper, we show that KRNL is effective for integrating multiple pieces of knowledge expressed in natural language by multiple persons. While the scale of experiment is still small, it successfully produces an integrated KRNL representation and then its corresponding natural-language sentence to be presented to a human user. Our approach of this paper demonstrates a way towards a new style of knowledge sharing.

### 1. はじめに

近年、Web 上から膨大な情報や知識を得ることのできるようになったが、その中でも特に多いのが自然言語で記された情報である。一つの文書から望む情報や知識を得ることができない場合には、複数の文書を組み合わせることによって知識を獲得することができる。しかし、これらの文書を統合するには、何らかの形式化された知識表現を導入し、それに従った規則で知識を表現する必要がある。また同様に、知識共有システムに対しても、同じことが言える。つまり、統合的な知識管理が可能な何らかの知識表現を用いる必要がある。知識を入力するユーザが統一された知識表現に従った知識入力を行うことで有効な知識共有が行われるのである。Web 文書の統合、知識共有システムのいずれの場合においても、複数ユーザによる知識共有をいかに実現するかが重要な鍵となる。そして知識を容易に組み合わせられる知識表現を用いることで、有効な情報支援を行うことができると考えられる。

知識を組み合わせるための表現として、従来、述語論理、命題論理などの論理表現が研究されてきた。一階述語論理では、いくつかの引数をとる述語記号を用い、正や負の述語を連言や選言などの論理記号で結合した整合論理式によって知識を表現する。そして、論理式で表現された知識を導出原理により組み合わせることで推論を行う。一階述語論理は、数学的基礎に立脚し、完全性と健全性を持つ演繹推論系が存在し、表現力も豊かであるため、自然言語と論理に着目した研究はいくつか行われている [石塚 96]。

しかし、自然言語文を論理に変換する場合、表現が一意に定まらないという問題がある [石塚 96]。例えば、下の例は「円高 (円の為替レートは高い)」を述語で表現したものだが、述語をどう設定するかや引数に対して任意性がある。

high\_exchange\_rate(円).  
 exchange\_rate(円, 高).  
 value(exchange\_rate, 円, 高).

A: 友部 博教, 名古屋大学情報科学研究科, 愛知県名古屋市千種区不老町, 052-789-5878, tomobe@nagao.nuie.nagoya-u.ac.jp

同じ知識でも表現が異なると、知識の組み合わせによる推論が困難となる。知識が複数ユーザによって記述されることを想定すると、さらに表現の任意性の問題は大きくなる。これに対し、自然言語から述語表現へ変換する際に、推論に必要な程度に任意性を制限した知識表現を介せば、この問題をある程度解消することができる。

また、自然言語には字面を読むことで理解できる規則がある。例えば、「太郎は犬である」と「犬はたまねぎを食べない」という自然言語で記述された知識があればこの2つを組み合わせると「太郎はたまねぎを食べない」というが考えられる。「太郎は犬である」という自然言語は包含関係を持っており、そこから「太郎」も「たまねぎを食べない」という性質を持つことがわかる。また、「テオティワカン人はテオティワカン遺跡を作る」という自然言語で記述された知識があれば、「テオティワカン遺跡はテオティワカン人に作られる」という能動態と受動態の変換を行うことができる。

本論文では、KRNL(Knowledge Representation for Natural Language) という知識表現を提案する。これは、自然言語文を述語論理に変換する際に、表現の任意性に一定の緩い制約を課すことによって、組み合わせによる推論に必要な述語記号の表現の一致を確保し、また自然言語に内在する規則をあらかじめ与えておくことで、自然言語文で表現された知識の利用を簡単にかつ効率的に行うものである。

### 2. KRNL の表現法と推論

KRNL では自然言語による知識をユーザが KRNL で表現しておく、KRNL は自動的に一階述語論理に変換され、Prolog により推論が可能となる。そしてその結果を KRNL もしくは自然言語文としてユーザに返す。これを 1 に示す。

以降、本章では自然言語文を KRNL に変換する手順ならびに KRNL を用いた推論について述べる。

#### 2.1 原子概念・分子概念

KRNL では、自然言語の語を原子概念と呼ぶ。同義語 (synset) を概念としてひとつのラベルを与えるなどの拡張も考えられるが、ここでは自然言語との対応の容易さを確保するために、単純に名詞、動詞などの語を原子概念としている。



図 1: KRNL による知識変換の全体像

また、複合語も自然言語との対応からそのまま原子概念として扱う。

次に、原子概念と、その修飾語をあわせたものを分子概念と呼ぶ。「ピラミッド (ETCMOD 月)」「研究する (OBJ 自動要約)」などが分子概念である。(ETCMOD 月) や (OBJ 自動要約) は修飾ラベルと呼ばれる。分子概念は、

原子概念 (修飾ラベル 1 原子概念 1) (修飾ラベル 2 ...) ...

と定義される。分子概念は、自然言語では名詞句もしくは動詞句にあたる 1 にそれぞれの修飾ラベルを挙げる。名詞句では、属性値関係を表す “ATOF” と修飾を表す “MOD” という修飾ラベルを用意する。“MOD” ラベルには、“ISAMOD”、“HASMOD”、“ETCMOD” という 3 つの種類があり、名詞の修飾の役割に応じて使い分ける。“ISAMOD” では、包含関係を持つ修飾を記述する。例えば、「犬の太郎」というように、「犬」と「太郎」に包含関係がある場合には、“ISAMOD” を用いて「太郎 (ISAMOD 犬)」と記述する。“HASMOD” では、所有や存在の関係を持つ修飾を記述する。例えば、「花子の本」というように、「花子」と「本」の間に所有の関係があれば、“HASMOD” を用いて「本 (HASMOD 花子)」と記述する。これ以外の修飾については “ETCMOD” を利用する。

動詞句では、動作の主体を表す “AGT” と動作の受手を表す “OBJ”、それ以外の ETC を用いる\*1。例えば、次の例では、「花子」「バラ」とも「あげる」の目的語であるため並列に記述している。

自然言語: 「花子にバラをあげる」

↓

KRNL: あげる (OBJ 花子) (OBJ バラ)

## 2.2 概念関係

KRNL では、分子概念間の関係を概念関係と呼ばれる関係で記述する。[ISA][HAS] などが概念関係であり、さまざまな概念間の関係から同値関係や上位-下位関係、全体-部分関係など、特に推論に有用で一般的なものを選んでいく。KRNL で用いる概念関係を 2 にあげる。

概念関係 [EQ] は交換律 (つまり A[EQ]B B[EQ]A) が成り立つが、それ以外では成り立たない。しかし、[ISA] や [HAS] に対しては、逆向き関係として [INC]、[POF] が対応している。[DO] と [PV] の関係は、ラベルの操作を伴うのでやや複雑である。

A [ISA] B      B [INC] A

\*1 これは Fillmore の考案した格フレームにおける格の種類 [Fillmore 66] の中で、動作の主体を表すものを “AGT”、動作の受手を表すものを “OBJ”、そして場所の起点や動作の手段を表すものなどについては、KRNL において組み合わせの推論には関わらないためすべて “ETC” としている。

表 1: 修飾ラベル

ラベル	意味	例文
名詞句を表すもの		
ISAMOD	ISA-MODifier: 修飾関係	犬の太郎 太郎 (ISAMOD 犬)
HASMOD	HAS-MODifier: 修飾関係	花子の本 本 (HASMOD 花子)
ETCMOD	ETC-MODifier: 修飾関係	赤色の車 車 (ETCMOD 赤色)
ATOF	ATribute OF 属性表現	花の色 色 (ATOF 花)
動詞句を表すもの		
AGT	AGenT 動作の主体	父に食べられる 食べる (AGT 父)
OBJ	OBject: 動作を受ける対象	りんごを食べる 食べる (OBJ りんご)
ETC	上記以外	

この [ISA] と [INC] のような関係を逆向であると呼ぶ。2 の逆向とはそのような意味を持つ。

[DO] と [PV] については、動詞句を表す分子概念を用いるので、他の概念関係では出現しない修飾ラベル (“AGT” や “OBJ”) が使われている。[DO] と [PV] は概念関係の逆行とは違うが対応関係を持っている。これも KRNL の組み合わせ推論において用いられる。

A [DO] B (OBJ C)

C [PV] B (AGT A)

この知識変換は、自然言語で記された知識において、能動態と受動態の相互変換を表現している。これは自然言語において暗黙の了解であるものを明示化したものである。

## 2.3 概念関係式

概念関係によって分子概念を結びつけたものを概念関係式という。概念関係式は次のように表される。

分子概念 1 [概念関係] 分子概念 2.

否定の場合には、次のように記述する。

分子概念 1 [概念関係](not) 分子概念 2

KRNL では概念関係式が知識のひとつの基本単位となる。

## 2.4 変数、論理演算子

KRNL では、変数や論理演算子を用いることができる。まず、複数の分子概念を AND や OR でつなげたものを分子概念として用いることができる。

## 2.5 KRNL の推論

KRNL では、概念縮合と呼ばれる概念関係の組み合わせによる推論を行う。例えば、

メキシコ [HAS] アカブルコ.

アカブルコ [HAS] ピーチ.

から、

メキシコ [HAS] ピーチ.

表 2: 概念関係

記号	逆向	意味	例
ISA(is-a)	INC	帰属関係	太郎 [ISA] 人間 (太郎は人間である)
INC(include)	ISA	帰属関係	動物 [INC] 犬 (動物は犬を含む)
HAS(has-a)	POF	所有関係	車 [HAS] タイヤ (車にはタイヤがある)
POF(part-of)	HAS	所有関係	手 [POF] 人間 (手は人間の一部だ)
VAL(VALue)	なし	属性値関係	色 (attr 車) [VAL] 赤い
EQ(EQual)	EQ	等価関係	合衆国 [EQ] アメリカ
DO(do)		動作関係	太郎 [DO] あげる (OBJ 花子)(OBJ バラ)
PV		動作関係	花子 [PV] あげる (AGT 太郎)(OBJ バラ)

概念関係式 1:           メキシコ [HAS] アカブルコ  
 概念関係式 2:           アカブルコ [HAS] ビーチ  
 結合:           メキシコ [HAS] アカブルコ [HAS] ビーチ  
 概念関係積:           [HAS] × [HAS] = [HAS]  
  
 結論:           メキシコ [HAS] ビーチ

x (MOD p) [DO]  
 y (OBJ a) (OBJ b) (ETC c) (ETC d).  
 ↓  
 do(x, p, y, a, b, c, d).

図 2: 概念縮合の例

を導く。これは、「A [HAS] B かつ B [HAS] C ならば A [HAS] C」より簡略化して [HAS] × [HAS]=[HAS] と書く。この関係をまとめたのが 3 である。[ISA] や [HAS] のもつ推移律や上位概念の性質の継承などをまとめたものである。

表 3: 概念関係間の組み合わせルール

	ISA	INC	HAS	POF	VAL	EQ	DO	PV
ISA	ISA		HAS					
INC		INC						
HAS	HAS		HAS					
POF		POF						
VAL				POF				
EQ					VAL			
DO	ISA	INC	HAS	POF	VAL	EQ	DO	PV
PV	DO	PV				DO	PV	

2.5 にこの概念縮合を用いた推論の例を示す。

### 3. KRNL の論理への変換と実装

#### 3.1 [ISA][INC][HAV][POF][EQ] の述語表現

これらの概念関係を述語とし、分子概念と修飾ラベルの値を引数として表現する。KRNL から述語論理への変換は次のようになる。

$$x \text{ (MOD } p) \text{ [cr]} y \text{ (MOD } q).$$

$$\Downarrow$$

述語論理:  $cr(x, p, y, q).$

MOD ラベルの値が無い場合には、nil を用いる。

#### 3.2 [DO] の述語表現

[DO] は左辺分子概念は [ISA] や [HAS] と同様であるが、右辺分子概念が他とは異なる。OBJ や ETC といったラベルは [DO] の右辺分子概念でのみ使われるので、述語論理への変換は次のようになる。

各ラベルの属性値を対応する引数場所に納めることによって記述する。[DO] で扱う OBJ ラベルと ETC ラベルはそれぞれ 2 つずつと制限する。

#### 3.3 概念関係積の記述

次に、KRNL の組み合わせによる推論である概念関係積を記述する。概念関係積は概念関係と概念関係の組み合わせで表現するので、述語論理へ変換するにはルールとして記述する。

3.31 二つの知識の概念関係が [ISA] [INC] [HAS] [POF] [EQ] である場合には、KRNL 「a [isa] b.」と「b [isa] c.」から「a [isa] c.」を導き出す概念関係積を述語論理のルールにすると、以下のようになる。

$$isa(x, p, y, q) \quad isa(y, q, z, r)$$

$$isa(x, p, z, r).$$

#### 3.4 Prolog による知識共有システムの実装

KRNL は複数ユーザによる知識入力を行う知識共有システムを想定している。そこで、KRNL の実証システムとして、KRNL 処理系および知識共有システムを、Prolog で実装した。このシステムでは、まず KRNL で記述された知識を Perl によって述語論理に変換し、述語論理の推論系として Prolog を用いた。現行の KRNL は時制等の扱いはせず、すべて現在形の知識を扱うことにしている。前節でも示したように、1 階述語ホーン節論理によって基礎付けを行っているので、実装には Prolog が適している。

##### 概念関係式の記述

KRNL の概念関係式を述語論理表現として変換する。例えば「メキシコには遺跡があります」という自然言語で記された知識は KRNL では

メキシコ [HAS] 遺跡.

となるが、述語論理では

has(メキシコ, nil, 遺跡, nil).

となる。

##### 概念関係積の記述

KRNL で記述された知識の組み合わせによる推論として概念関係積を利用する概念関係 [ISA] や [EQ] は探索の際無限に

展開されてしまう可能性がある。そこで、探索の深さを限定することによってこの問題を解決している。次にあげるのは、[HAS] と [ISA] から [HAS] になるという概念関係積を述語論理に直したものである。

KRNL	X (MOD P) [HAS] Y (MOD Q),
	Y (MOD Q) [ISA] Z (MOD R)
	X (MOD P) [HAS] Z (MOD R).
述語論理	has(X,nil,Z,nil):-
	has(X,nil,Y,nil),
	isa(Y,nil,Z,nil).

なお実際には再帰的な計算になるので、探索の深さを制御している。<sup>\*2</sup>

## 4. 議論

### 4.1 自然言語との対応

自然言語から KRNL の変換

自然言語文で記された知識の種類は、大きく分けて動作を表すものと状態を表すものに分かれる。例えば、「太郎が花子に花を贈る」というのは動作を表す自然言語文であるし、「友部博教は東京大学の学生である」というのは状態を表す自然言語文である。

状態を表す自然言語文の知識は、文中に二つのフレーズが出現し、その関係によってさらにいくつかに分けることができる。「友部博教は東京大学の学生である」というのは、「友部博教」と「東京大学の学生」という二つのフレーズが包含関係であり、「車にはタイヤがある」というのは「車」と「タイヤ」が所有関係であり、前者とは違う関係として位置づけることができる。

KRNL の概念関係は、この自然言語文で記された知識の種類に対応している。自然言語文で記された知識が「動作」であれば概念関係 [DO] を、「状態」を表すものであれば、包含関係なら [ISA] を、所有関係なら [HAS] を用いる。これは、自然言語文における知識の種類に対応しているため、知識が任意性なく表現できることになる。

KRNL から自然言語への変換

KRNL から自然言語へ変換するには、前節で述べた自然言語文で記された知識の種類を用いれば良い。自然言語で印された知識は、大きく分類すると状態と動作に二分される [中川 96]。[DO] であれば動作を表す表現に、[ISA] であれば包含関係の状態を表す表現である。例えば「友部博教 [ISA] 学生」であれば、「友部博教は学生である」となる。

また、修飾ラベルを用いて分子概念を表現しているが、分子概念では係り受けも考慮して記述している。また、修飾ラベルに対応する自然言語も一意であるので、KRNL から自然言語の変換は容易に行うことができる。

### 4.2 述語論理との対応

4 節で示した通り、KRNL から述語ホーン節論理への変換は一意に対応させることができる。よって、KRNL は述語ホーン節論理の持つ性質（推論の健全性や、完全性）を継承することができる。推論としては頑健なものとなる。

一方、すべての述語ホーン節論理表現を KRNL で表現することはできない。そのため、KRNL は述語ホーン節論理のサブセットとして位置づけることができる。

<sup>\*2</sup> 実際には、探索の深さと深さの限度を引数として追加した述語を用いる。isa では isa-rule(Limit,Depth,X,P,Y,Q).(Limit は深さの限度、Depth は現在の深さ) を用いる

### 4.3 類義語、同義語や背景知識に対する対応

例えば「赤色」と「赤い」のように、同義語に関しては表記ゆれが生じることがある。また、専門用語を扱うようなデータベースであれば、例えば「自動要約は自然言語処理に含まれる」といったような背景知識も推論処理に必要なになるだろう。KRNL では単純な字面のマッチングしか行わず、ルール記述にもユーザに必要な最低限のものしか要求しない。そのため、類義語、同義語や背景知識に対応するには、語彙体系や EDR 電子化辞書のような外部知識源が必要となる。これは KRNL のシステム外の問題であると考えられる。こういった外部知識源を利用する場合には、KRNL では例えば「赤色」と「赤い」の場合には「赤色 [EQ] 赤い」と記述することもできる。

## 5. おわりに

自然言語で記された大量の情報/知識の統合的高次利用を図るアプローチの一つの基礎となると考える、自然言語表現と整合性が良い知識推論管理システム KRNL (Knowledge Representation for Natural Language) について述べた。そして複数の人によって自然言語文により記された断片的な知識を KRNL に変換し、効果的に統合する知識共有の新しい方法を示した。自然言語で記された情報/知識は増大していくので、今後、KRNL の基盤の充実・確立と、多様な応用機能の研究開発を進めていく予定である。

複数人により記述された知識には相矛盾するものも含まれる可能性があるが、各人の記述知識をワールドとして管理し、ワールド内では無矛盾として、矛盾しない範囲で複数ワールドを組み合わせる知識管理なども知識共有のためには必要になると考えている。また、人により知識の信頼度は異なることから、上記ワールドに信頼度を付し、最も高い信頼度のワールドの組み合わせによって得られる推論結果を優先する等のメカニズムも必要になると考えている。

## 参考文献

- [石塚 96] 石塚 満：知識の表現と高速推論, 丸善 (1996)
- [Fillmore 66] Fillmore, C. J.: Toward a modern theory of case, pp. 361-375 (1966)
- [Iwanska 00] Iwanska, L. and Shapiro, S. C.: *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*, AAAI Press/MIT Press (2000)
- [中川 96] 中川 裕志：順接複文における主語の共参照関係の分析, 自然言語処理, Vol. 3, No. 2, pp. 59-74 (1996)