

概略推論を用いた質問応答システムにおけるゴール抽出法

Extracting a Goal Sentence for Answering Mechanism using ARSK

佐藤 雅彦^{Ⓐ1} 石川 勉^{Ⓐ1} 河岡 司^{Ⓐ2}
 Masahiko Satou Tsutomu Ishikawa Tsukasa Kawaoka

^{Ⓐ1} 拓殖大学 工学部 情報工学科
 Department of Computer Science, Takushoku University

^{Ⓐ2} 同志社大学 工学部 知識工学科
 Department of Knowledge Engineering and Computer Sciences, Doshisha University

We proposed about " ARSK (Approximate Reasoning with Similar Knowledge) " which could obtain an approximate solution even if the knowledge stored in the problem solving system was incomplete. ARSK operates based on SLD, which is a backward reasoning scheme. So, it can obtain solutions only when a question about a conclusion is asked on condition that the causes concerning to the conclusion are given. That is, ARSK can't get any solutions in the reverse case. This paper proposes a method to select a goal from knowledge given in a question, to get possible solutions even in such a case.

1. はじめに

我々は、対話型ロボットへの適用を想定した自由対話システムについて研究している [1]。このシステムではその質問応答部に、これまで提案してきた概略推論法 " ARSK (Approximate Reasoning with Similar Knowledge) " を用いている [2]。ARSK は、知識が不完全な場合でも概略的な解を導く推論法である。具体的には、問題知識の一部が不足している場合には、概念ベース [3] を利用して問題知識の中から類似する知識を検出し、それを用いて概略的に推論を行う。ここでの推論は後ろ向き推論の SLD 法をベースに、2つの素式が類似すればそれを用いて導出を行うことにより実行される。

以上で述べたように、ARSK は SLD 法を用いているため、ゴールとなる知識は結論に近い形態で与えられることを前提としている。すなわち、原因側の知識が与えられた上で、結論側の質問 (ゴールとなる) がなされることを前提とする。従って、質問によっては充分な知識があるにもかかわらず回答できない場合がある (上記と逆の場合)。本稿では、このような場合にも可能性のある解を提示するための、ゴールとすべき知識の抽出法について提案する。

2. 概略推論法 " ARSK "

2.1 知識表現と推論方法

ARSK では、知識は基本的には 1 階述語論理で表現され [4]、素式のフォーマットは下記のような形式としている。

述語 (ラベル $_1$: 引数 $_1, \dots$, ラベル $_n$: 引数 $_n, \dots$)

ここで、述語と引数は単語あるいはその結合により表される。また、述語は動詞、名詞、形容詞であり、動詞の場合には引数はフィルモアの格文法の深層格を用いる。知識表現法の詳細については、文献 [5] を参照されたい。

ARSK は、以上のように表現される知識に対して概略的な推論 (導出原理に基づく) を行うが、これは類似知識の適用によりなされる。具体的には、従来の導出がルール左辺ともう一

方のルール右辺が同一ならば、二つの節から新しい節 (導出節) を生成するのに対し、ある程度類似していれば導出節を生成させる (類似導出)。類似導出のイメージを図 1 に示す。

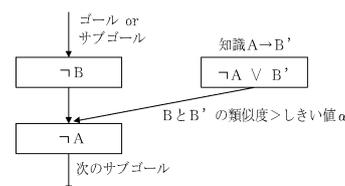


図 1 類似導出のイメージ

2.2 従来の ARSK の問題点

従来の ARSK の問題点について例をあげて説明する。以下のような質問がなされたとする。

" 太郎は車の構造について知識があります。太郎は車を直せますか "

この質問は述語論理で表現すると下記のようなになる。

【事実知識】

知識 (agt:太郎, obj:車・構造)

【質問知識】

直す (agt:太郎, obj:車)

これに対し、問題知識として、以下が用意されていたとする。

【問題知識】

知識 (agt:x, obj:車・構造) 修理 (agt:x, obj:車)

この場合、質問知識は事実知識より結論に近い形態となるので、質問知識をそのままゴールとする従来の ARSK でも質問に回答することができる (図 2.a)。

次に、同様の状況において次の質問がなされた場合について考える。

“ 太郎は車を直せます。太郎は車についての知識がありますか ”

これは質問が原因を問う場合である。この場合、前述の場合に対し、質問知識と事実知識が逆となり、質問知識をそのままゴールとして用いると、これと照合するルールがないため解の導出が不可能となり質問に答えることができない(図 2.b)。



図 2.a 回答可能な場合

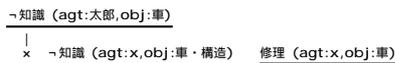


図 2.b 回答できない場合

4. ゴール抽出法の具体化

以上の方法を質問応答システムにおける前処理と位置付け、ゴール抽出モジュールとして具体化した。以下、これについて具体例を用いて説明する。次のような質問がなされ、それに対し図 4 に示した問題知識が与えられているとする。このモジュールでの処理は以下の (1)~(5) の手順で行われる。

“ 太郎は車の整備に熟練していて、車を修復できます。太郎は何について勉強しましたか。 ”

学習 (agt:x,obj:車・構造) 知識 (agt:x,obj:車・構造)
 知識 (agt:x,obj:自動車) 訓練 (agt:x,obj:車・整備) 合格 (agt:x,obj:整備・試験)
 合格 (agt:x,obj:整備・試験) 修理 (agt:x,obj:車)
 趣味 (agt:x,obj:車) 運転 (agt:x,obj:車)

図 4 問題知識

3. ゴールとなる知識抽出の考え方

本稿では、質問が原因あるいは結論のいずれに近い形態で問われても回答可能とするため、質問で得た知識の中で最も結論に近い知識を抽出し、ゴールとする方法について提案する。

この方法では、あらかじめすべての問題知識を木構造化しておく。具体的には、ある 2 つのルールのうち、一方の左辺(原因)ともう一方の右辺(結果)が類似すればつなげ、この操作を繰り返すことにより問題知識をすべて木構造化する。知識は“(原因) (結果)”の形式であるので、木の下方向ほど結論に近いといえる。次にその木の中の知識と類似する、質問で得た知識を対応させる。木の下方向の知識と対応した知識は、その質問で得た知識の中でも結論に近いとみなし、その知識をゴールとして抽出する。

【問題】 太郎は薬を飲みました。太郎は風邪なのでしょうか？

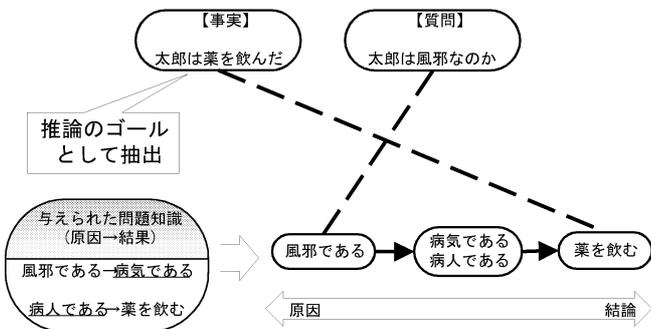


図 3 ゴール抽出法のイメージ

以下、図 3 に沿って具体的に説明する。図 3 では“ 太郎は風邪なのだろうか ”という原因についての質問をしている。この場合、この質問に関する与えられた 2 つの問題知識のうち、“ 病気である ”と“ 病人である ”が類似するならば、その 2 つをつなげる。よって、与えられた問題知識の中では、“ 薬を飲む ”という知識が結論となる。図 3 において、木に含まれる“ 風邪である(原因) ”という知識は質問知識に対応し、同じ木の下方向にある“ 薬を飲む(結論) ”という知識は“ 太郎は薬を飲んだ ”という事実知識に対応したとすれば、その事実知識をゴールとして抽出する。

なお、この方法は結果から原因を求めるものであり、当然のことながら得られた解が正しいとはいえないことに注意する必要がある。

(1) 知識のリスト化

初めに、与えられた知識のルールの引数を除いて概念ごとに分割する。そして、各概念に番号を付与し、その概念の上方(前)と下方(次)の概念の番号を示したリストを作成する。そのリストを表 1 に示す。図 4 の初めのルールの“ 学習 知識 ”でいうと、“ 学習 ”に番号 1、“ 知識 ”に番号 2 をふる。番号 1 の“ 学習 ”は次の概念番号は 2 の“ 知識 ”を指し、番号 2 の“ 知識 ”については前の概念番号は 1 の“ 学習 ”となる。

表 2 知識のリスト

番号	概念	次の概念の番号	前の概念の番号
1	学習	2	
2	知識		1
3	知識	5	
4	訓練	5	
5	合格		3,4
6	合格	7	
7	修理		6
8	趣味	9	
9	運転		10

(2) 知識の連結による木構造化

次に、元のルールの左辺と右辺が類似すれば、その 2 つの概念の指す次の概念番号と前の概念番号を統一する。例えば表 2 において、番号 2 の“ 知識 ”と番号 3 の“ 知識 ”が類似(この場合は同一)したとすれば、番号 2 の“ 知識 ”の次の概念番号に番号 5 を加え、番号 3 の“ 知識 ”の前の概念番号に番号 1 を加える。この操作によりルールを繋げて仮想の木構造を作る。操作後のリストを表 3 に示す。

表 3 知識の連結後のリスト

番号	概念	次の概念の番号	前の概念の番号
1	学習	2,3	
2	知識	5	1
3	知識	5	1
4	訓練	5	
5	合格	7	2,3,4
6	合格	7	2,3,4
7	修理		5,6
8	趣味	9	
9	運転		10

(3) 木構造化された知識と問題知識との対応

次に、質問で得た知識の概念を、リスト中の概念に対応させる。この操作により、質問知識に対して事実知識は結論または原因のどちら側にあたるのかがわかる。対応後のリストの木構造を図 5 に示す。

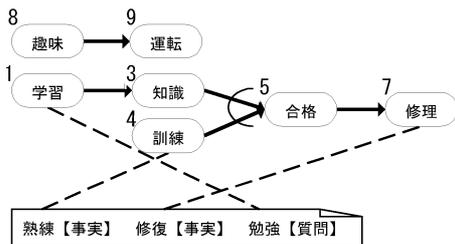


図 5 知識の木構造と入力知識の対応

(4) 不必要な知識の除去

これまでの操作で作成された木の中で、どのルールとも繋がっていない非連結の知識は除去する。この問題の場合、趣味 (A:x,O:車) 運転 (A:x,O:車) はどの対話で得た知識とも対応せず非連結であるため除去される。

(5) ゴールの抽出と信頼度の付与

質問知識と対応した知識から木の下方へ進み、事実知識と対応した知識を探索する。この例の場合、質問知識“勉強”と対応した“学習”という知識から下方へ探索すると、事実知識“修復”と対応した“修理”という知識が見つかり、この事実知識“修復”をゴールとして抽出する。また、複数の知識が抽出された場合の選択の基準として、抽出された知識に信頼度を付与する。信頼度は質問知識と対応した知識から、抽出されたゴール知識と対応した知識までの経路で行われた、知識の対応や連結部分における類似度の積とする。

5. システムの動作確認

職業に関する 215 個の知識からなる知識ベースを作成し、システム全体についての動作確認を行った。例えば 4 章の例で抽出された事実知識“修復”をゴールとして推論を行うと、図 6 のような導出過程が得られ、質問の回答となる解が導かれる。すなわち、最後の導出について x と単一化され、“車・構造”が解として得られることになる。

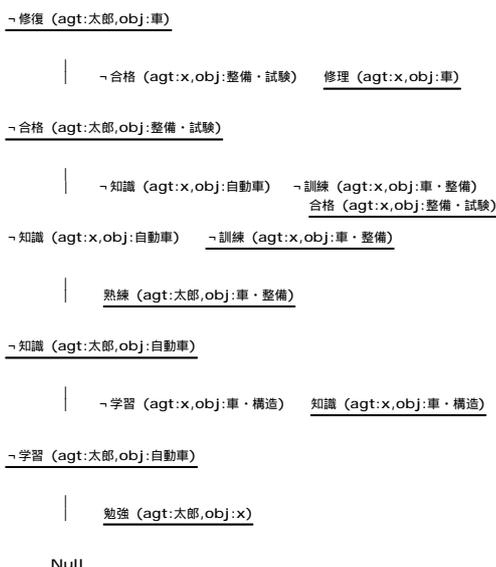


図 6 解“車・構造”の導出過程

6. まとめ

結果が与えられ原因を問う質問に対し、可能性のある解を得るための、ゴール知識の抽出モジュールについて提案した。また質問応答システム前処理としてモジュール化し、その正常動作を確認した。

参考文献

[1] 小林 高志: “感情表現と質問応答能力を備えた知的自由対話システム”, 人工知能学会研究会資料 SIG-SLUD-A003-1 pp.1-8

[2] Nguyen Viet Ha 他: “知識の類似性を利用した概略推論法の研究”, 電子情報通信学会論文誌 D-I Vol.J84-D-I No.4

[3] Nguyen Viet Ha 他: “単語の意味の類似性判別のための大規模概念ベース”, 情報処理学会論文誌 vol.43 No.10 pp.3127-3136

[4] 荒屋真二: “人工知能概論”, 共立出版株式会社

[5] 石川勉: “言葉で考えるコンピュータの実現にむけて”, 2003 年度 人工知能学会全国大会 (第 17 回) 論文 3C1-04