

意味グラフ照合による 質問応答システム Metis の回答精度向上

- 質問文解析フェーズと検索フェーズの改良を中心として -

西岡 晋太郎[†] 久保田 裕章[†]
坂東 晃文^{††} 原田 実^{†††}

我々は、2003年から意味グラフベースで質問文と知識文を照合することで回答抽出を行う質問応答システム Metis を開発している。本研究では、factoid 型質問に対する回答精度向上を行った。質問文における質問箇所と同定とそれに対する意味制約の付与方法の改善などによるグラフ照合の精度向上、知識文のデータベースにおいて深層格を伴うインデキシングと役割や組織名を含むキーワード抽出などの改善を行った。これにより、NTCIR-6 CLQA のデータに基づく評価実験を行った結果 79% の回答抽出精度を得た。

Answering Improvement of QA System Metis Based on Semantic Graph Matching

- Focused on Question Analysis and Knowledge Retrieval -

Shintaro Nishioka[†] Hiroaki Kubota[†]
Akifumi Bando^{††} and Minoru Harada^{†††}

Since 2003, we have been developing a question answering system “Metis” which extracts the answer by matching the question sentence with the knowledge sentence under the semantic graph framework. In this research, we focused on the factoid type questions. Actually, the accuracy of graph matching was improved by the identification of question parts in question sentences and by giving the meaning restriction to the question parts. Besides, introducing the index with a deep case for a verb in the database of knowledge sentences and the improvement of key word extraction including roles and organization names were done. As a result, the answer extraction precision of 79% was obtained in the evaluation experiment based on the data of NTCIR-6 CLQA.

[†] 青山学院大学大学院理工学研究科理工学専攻知能情報コース

^{††} 日本ユニシス株式会社

^{†††} 青山学院大学理工学部情報テクノロジー学科

1. はじめに

近年、自然文で書かれた質問文に対して回答を抽出する質問応答システムの研究が盛んに行われている。先行研究では、質問文と知識文をある基底表現に言い換えてから照合を行うシステム[1]、質問文が要求している回答のタイプを詳細に分類するシステム[2]、質問文と知識文の係り受け構造の類似性から回答候補を得るシステム[3]、機械学習を用いた学習型質問応答システム[4]、統計的機械翻訳を利用した言語横断質問応答システム[5]などが提案されている。[1]では、質問文と知識文をうまく照合できるように、すべての文を基底表現に言い換えている。この手法では、質問文と知識文を確実に照合し解を得るため確実性は高いが、どの表現が基底表現であるかを決定し膨大な変形規則を用意しておかなければならない。[2]では、質問文の疑問詞に意味的カテゴリ（意味制約）を設定し、解を大雑把に推定している。現状でも簡単に利用できる手法だが、疑問詞の意味制約のみでは照合の確実性が乏しいという問題がある。[3]では、質問文と知識文の単語の出現頻度と質問箇所への意味制約だけでなく、文間の係り受け構造の類似性までを評価しノード間の最短路に基づく得点を計算し回答順位付けを行っている。[4]では、日本語による質問応答システムの主要なコンポーネントである質問文解析と回答抽出処理を SVM による機械学習で構築している。約 15000 の質問に対し評価実験を行ったところ、TOP5 までの正解率が 55.7% であり MRR 値が 0.393 であったと報告されている。しかし、SVM を利用したシステムでは、訓練データを作成する際に大量のデータを学習器への入力形式に変換すること（例えばタグ付けなど）や正解事例を人間が指定する必要がある。[5]では、統計翻訳モデルを質問応答プロセスに組み込むことで言語横断質問応答を実現しており、NTCIR-6 CLQA[6]においてはコンテスト参加者中最高位の 59.5% の精度を示している。

原田研究室では、語の意味と語間の深層格を評価できる意味解析を用いれば現在提案されているシステム以上に高精度な質問応答システムが開発できることに着目し、2003年から意味解析システム SAGE[7,8,9,10,11]を用いて意味グラフベースで質問文と知識文を照合することで回答抽出を行う質問応答システム Metis[12]を開発している。2006年度に開発した Metis2006[13]では、NTCIR-6 の CLQA と QAC[14]に参加したが、プロトタイプであったため回答抽出精度は他システムと比べて中程度のものであった。

本研究の目的は、従来の Metis を質問文解析フェーズと検索フェーズを中心に改良して、factoid 型質問に対する回答の精度を向上することである。

2. 質問応答システム Metis の概要

Metis は、図 1 に示すように、まず自然言語で入力された質問文と新聞記事や Web 中の文章（知識文）の両方を SAGE で意味グラフに変換する。ここでは、文は文節を頂点、係り受けを辺にしたグラフで表現され、文節の主辞には EDR 辞書[15]の語意が、辺には SAGE が定めた 36 種の深層格のどれかが割りあてられる。この両グラフの共通部分グラフ（閾値以上の語意類似度を持つ頂点ペアとその間の閾値以上の格類似度を持つ辺ペア）の大きさで文間の類似度（グラフ類似度）を判定する。例えば、図 1 に示すように Metis に質問文「ペスト菌を発見した細菌学者は誰ですか。」が入力されると意味解析を行った後、質問文解析を行う。この処理では疑問詞が特定され、それに与える意味制約（誰）を決定する。次に、質問文から検索エンジン呼び出すためのキーワード（ペスト菌、発見、細菌学者）を抽出し知識文検索を行う。得られた知識文と質問文の意味グラフを照合しグラフ類似度が最上位の知識文から回答（北里柴三郎）を抽出する。補助機能としてそれ以外の回答もグラフ類似度順に表示する。

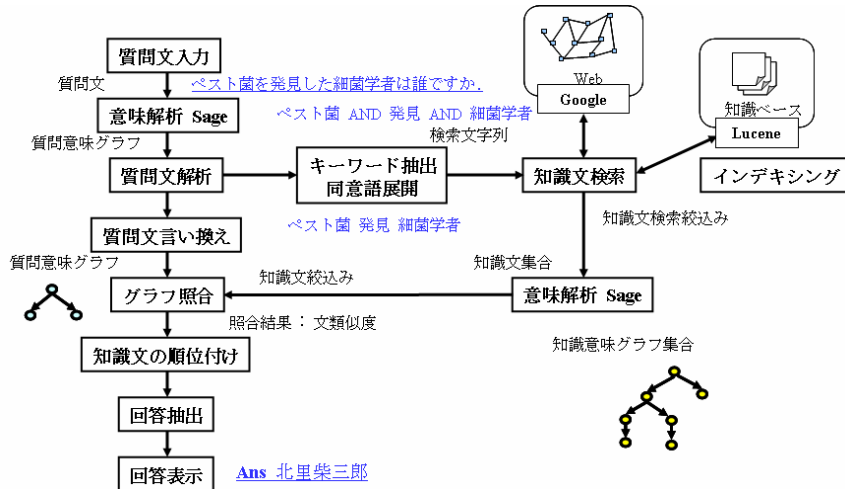


図 1 Metis システム構成

① 質問文解析

まず、Metis に入力された質問文は SAGE により意味解析され質問グラフに変換される。質問グラフ中に疑問詞があれば質問箇所特定に進む。質問箇所特定処理は質問が何を回答として求めているのかを判別し、factoid 型質問（人物、時間といった

単語を回答とする質問）の場合には質問グラフ中において、疑問詞を持つノード（質問ノードと呼ぶ）に質問箇所であることを示すフラグを立てる。質問ノードを特定した後、factoid 型質問の場合に質問箇所がどういった概念であるべきかを表わす意味制約（EDR 辞書における概念 ID）を与える。

② 知識文検索

知識文検索の流れを図 2 に示す。まず、質問文解析が終了した質問グラフに対してキーワードを抽出する。キーワードを抽出する際には、Referent キーワード（リファレント（固有表現）を持つノードのこと）、Must キーワード（共起関係子として主題助詞“は”を持つノードと、“最”、“初”、“一番”を持つノードのこと）、Normal キーワード（Must キーワード以外のノードで、EDR 詳細品詞として“普通名詞”、“サ変名詞”、“数詞”、“時詞”、“単位”、“後置助数詞”、“動詞”、“形容詞”、“形容動詞”を持つノードのこと）の 3 タイプに分けて抽出する。例えば「世界的名画「モナリザ」の作者は誰ですか。」という質問文の場合、Referent キーワードとして“モナリザ”、Must キーワードとして“名画”と“作者”、Normal キーワードとして“世界的”が抽出される。生成したキーワードの組合せから検索エンジンの仕様にあった AND 検索用の検索文字列を作成し、検索エンジン呼び出す。検索された知識文を SAGE で意味解析して知識グラフを獲得し、その中から有効な知識を選び出しグラフ照合へと進む。

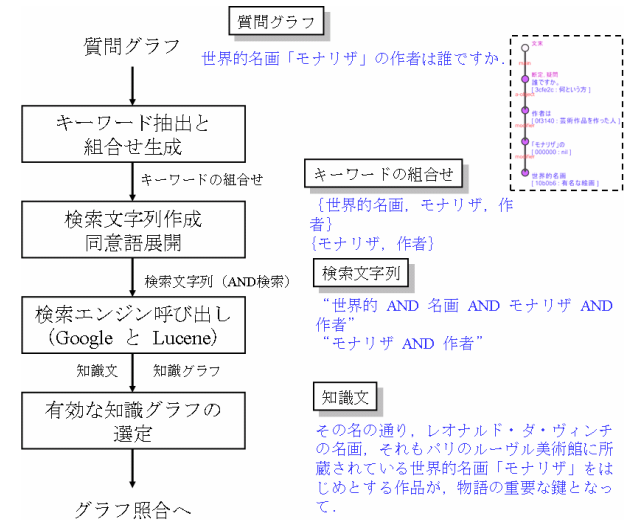


図 2 知識文検索の流れ

③ グラフ照合

検索で得られた知識グラフと質問グラフとの類似性を調べるためにグラフ照合を行う。グラフ照合は、はじめに質問グラフのノードと知識グラフのノードそれぞれの概念を C_1 , C_2 としその両ノード間の概念類似度 (式 1) を求める。尚、式 1 で 2 つの概念 C_1 と C_2 の概念類似度は EDR の概念体系木上での共通上位概念 $c(c_1, c_2)$ までの距離に基づいて計算する。この概念類似度とモダリティ得点 (式 3) と文節不一致係数とをかけたものをノード類似度 (式 4) とする。ただし、両ノードが固有表現を表す場合は表記による文字列類似度 (式 2) を基に計算する。また、モダリティとは、文が指す内容に対する話し手の判断や心的態度を表し、式 3 に示すように 4 種のモダリティの詳細種別ごとにあらかじめ与えた得点の積として計算する。

このノード類似度が閾値を超えたものを照合ノードペアとする。すべての照合ノードペアに対するノード類似度を合計してノードグラフ類似度 (式 5) とする。

$$\text{概念類似度} = \frac{2 \times d(c(c_1, c_2))}{d(c_1) + d(c_2)} \quad d(c): \text{概念}c\text{の深さ} \quad (1)$$

$$\text{文字列類似度} = \frac{2 \times \text{colen}(s, t)}{\text{len}(s) + \text{len}(t)} \quad \begin{array}{l} \text{len}(s): \text{文字列}s\text{の長さを返す} \\ \text{colen}(s, t): \text{文字列}s\text{と文字列}t\text{の共通文字列の長さを返す} \end{array} \quad (2)$$

$$\text{モダリティ得点} = \text{真偽判断} \times \text{価値判断} \times \text{発話} \times \text{程度の各グループ得点} \quad (3)$$

$$\text{ノード類似度} = (\text{概念類似度} \text{ or } \text{文字列類似度}) \times (\text{モダリティ得点}) \times \text{分節品詞不一致係数} \quad (4)$$

$$\text{ノードグラフ類似度} = \frac{\sum(\text{ノード類似度})}{\text{質問グラフのノード数}} \times 50 \quad (5)$$

一方、このような質問グラフノード間のアークとその両端のノードに対応する知識グラフノード間のアークとのアーク類似度を、それぞれのアークが持つ深層格が表 1 に示す深層格の類似グループのどれに共に属しているかで最右欄の値として定める。このアーク類似度を照合する全アークペアに対して合計しアークグラフ類似度 (式 6) とする。

$$\text{アークグラフ類似度} = \frac{\sum(\text{アーク類似度})}{\text{質問グラフのアーク数}} \times 50 \quad (6)$$

表 1 深層格グループ

グループ名	属する深層格名	アーク類似度
動作・陳述の主体	agent, o-agent, a-object, object, scene, modifier	0.90
時制系列	time, time-from, time-to, duration, sequence, reverse, cooccurrence, manner, modifier	0.90
動作の対象	object, goal, implement, material, source, o-agent, basis, beneficiary, modifier	0.85
修飾表現	a-object, modifier, possessor, manner	0.90
理由・原因	cause, condition, manner	0.80
動作の目標	goal, beneficiary, purpose, manner, modifier	0.85
場所	place, goal, from-to, location, scene, source, manner, modifier	0.90

ノードグラフ類似度とアークグラフ類似度の計算後、その和をグラフ類似度 (式 7) とする。

$$\text{グラフ類似度} = \text{ノードグラフ類似度} + \text{アークグラフ類似度} \quad (7)$$

factoid 型の質問の場合、“誰”や“どこ”といった質問ノードに対しては概念類似度を計算する際に表 2 最右欄に示すような概念 ID を追加する。これを意味制約という。例えば、疑問詞タイプが“誰”であるならば“人名”といった質問内容に則した意味制約を用いる。これにより factoid 型の質問の場合、質問ノードが回答ノードと照合するようになる。グラフ照合においては、通常ノードが一つの概念 ID からなる意味制約を持つものに対して、質問ノードに対しては複数の意味制約を持つことを許し、照合時においては式 1 の概念類似度の計算において複数の C_2 に対する計算結果の平均を概念類似度とする。

④ 回答抽出

Metis はグラフ照合の後、factoid 型質問の場合は質問ノードと照合した知識文のノード (回答ノード) を回答として返す。また、回答ノードへの入力辺または回答ノードからの出力辺の深層格が“modifier”または“a-object”で、EDR 詳細品詞が“普通名詞”、“固有名詞”または“未定義語”であるノードが存在する場合は回答を纏め上げる処理を行う。

3. 従来の Metis の問題点

以下に、具体的な3つの工程における問題点を挙げる。

① グラフ照合

質問文において質問ノード A と問われている内容を表すノード B が別の文節として分離している場合に、ノード A と照合すべき知識文のノード C がノード B に照合し正しい回答が返らない。

② 知識文検索

「黒澤明監督」のように役職や組織名等を持つものに対して、「黒澤 AND 明 AND 監督」のような検索文字列では正答を得るための知識文を抽出しにくい。また、カタカナからなる形態素の表記のブレに対応していない。

③ 回答抽出

構造まで一致しなくともキーワードを多く含む知識文には正解が含まれていることが多いが、このことが回答順位に反映されない。

4. グラフ照合における改良

先に述べたグラフ照合における問題点は、質問ノードと問われている内容を表すノード（実体ノードという）が分離している場合に特に顕著となる。そこで質問文解析において両者のノードを特定し、これらが別文節として存在する場合は実体ノードを質問ノードに融合する処理を行う。

まず、質問ノードと実体ノードが別文節かどうかを表す質問タイプ、疑問詞で問われている対象の種別を表す疑問詞タイプ、質問ノードに追加する意味制約の関係を表2にまとめた。ここで、質問タイプが独立型とは「二〇〇〇年のオリンピックはどこで開催されましたか。」のように、質問ノード（どこで）のみが存在し、実体ノードがない場合である。また、質問タイプが冗長型とは「S80 はどのメーカーの車ですか。」のように、質問ノード（どの）と実体ノード（メーカーの）が別文節として存在する場合である。さらに NANIX 型とは「黒沢明監督は生涯で何本の映画を作りましたか。」のように、疑問詞（何）と実体ノード（本）が別形態素だが一文節内に存在する場合である。

表 2 Factoid 型質問に対する質問文の分類

質問タイプ	疑問詞タイプ	疑問詞	質問ノードに疑問詞が持つ概念IDに加えて追加する概念ID（意味制約）
factoid 型 (独立型)	人	誰, だれ, どなた, どちら, どちら	3f961c 人間の具体的な名前
			30f6b0 人間
			0f15aa 役割で捉えた人間
			30f6b7 組織内の役割で捉えた人間
			0e3252 彼
	時	何時, いつ, 何時頃, 何時ごろ, いつ頃, いつごろ	444878 人間の一般的な名前
			444877 実在した特定の個人をさす名前
			30f776 時
			3f9882 時間点
			30f7aa 数量
	もの	何, なに, なん, 何と, なんと, 何という, 何と言う, どの, どれ, どちら	30f800 出来事
			30f93d 計量の単位
			30f7e4 事象
			3d017c 物事
			30f6ae 具体物
			444d86 抽象物
			3aa938 場所
			3aa912 自立活動体
			3aa911 人間または人間と似た振る舞いをする主体
			3aa963 状態
場所	何処, どこ, どの, どちら, どちら	3f965a 地域	
		444a40 国名で捉えた国家	
		3f9659 行政単位に対応する地名	
		3aa932 建造物	
		3aa938 場所	
		444a5f 地名で示す地域	
		30f769 教育のための建物	
		444b33 社会的要因からみた地域	
量	どれ位, どれくらい, どの位, どのくらい, どの程度, どのていど, 何度, なんど, いくら, いくつ	30f7aa 数量	
		30f93d 計量の単位	
		4446bd 金銭の単位	
		3aa963 状態	
		3c0fa9 パーセントという比率	
factoid 型 (冗長型)	冗長	実体ノードを修飾している疑問詞	実体ノードの主辞の概念ID
NANIX 型	何X	何+普通名詞, 後置助数詞, 単位	質問ノードの主辞の概念ID

Metis2008 における質問文解析の処理の流れを図 3 に示す。質問ノードが存在した場合は、質問タイプの判定を行う。「黒沢明氏が亡くなった月は。」のように、質問ノードが存在しない場合は実体ノードの主辞（この場合は“月”）の概念 ID と表 2 最右欄に示す疑問詞タイプ別の意味制約の概念 ID との概念類似度が最も高いものを疑問詞として文末に追加する（この場合は“いつ”が疑問詞として追加される）。疑問詞が追加できない場合（「野球はスポーツですか。」のような場合）は、Yes/No 文の判定を行う。Yes/No 文の回答の判定は、検索で得られた知識文集合のうちグラフ類似度が閾値を超える知識文が含まれる割合によって決定する。回答は 3 種類に分類され、閾値 0.8 以上ならば“はい、そうです”，閾値 0.5 以上ならば“たぶん、そうです”，閾値 0.5 未満ならば“違います”と回答が分類される。

質問タイプが factoid 型の場合、図 3 中の中段に示すように疑問詞タイプを判別し、疑問詞タイプに応じて意味制約を決定する。さらに、質問ノードと実体ノードが分離しているかを判定し、分離している場合は冗長型と判定する。

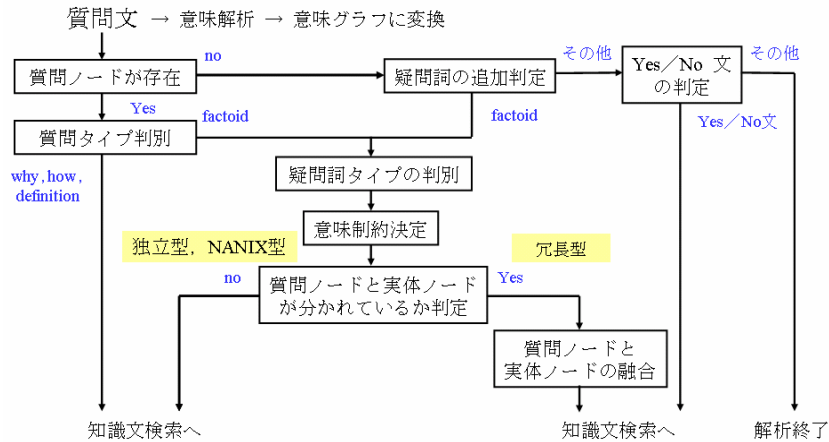


図 3 質問文解析の流れ

質問文解析において質問文が独立型や NANIX 型と判定された場合、質問ノードと照合した知識グラフ中のノード（回答ノード）を答えとして返す。例えば、独立型と判定される「二〇〇〇年のオリンピックはどこで開催されましたか。」に対して、検索された知識文章内に「2000年にシドニー五輪を開催するオーストラリアで、「観光新時代の幕開けにしよう」と取り組みが始まった。」という知識文があったとする。質問ノード（どこで）には表 2 に示すように意味制約として疑問詞タイプ“場所”の 7

つの概念 ID が与えられるので、これらの概念とのノード類似度が最も高い知識文のノード（この場合は“オーストラリアで、”）が回答ノードとなる。同様に、NANIX 型の場合も質問ノードとのノード類似度が高い知識文のノードを回答ノードとする。ただし、NANIX 型の意味制約には質問ノードの主辞の概念 ID を与えている。これにより、「黒沢明監督は生涯で何本の映画を作りましたか。」に対して、検索された知識文章内に「故・黒沢明「七人の侍」など傑作 30 本の監督を通じ、日本映画の名を海外にまで高め、世界のクロサワと呼ばれた功績に対して。」という知識文があった場合、質問ノード（何本の）の意味制約 (1e862b: 細長いものを数える語) と正解の回答ノード (30本の) の概念 ID (1e862b) が同じで類似度 1.0 となるので正しく照合する。

一方、質問文が冗長型と判定された場合、グラフ照合において質問ノードと実体ノードを融合して質問ノードとし、グラフ照合を行う。この場合、融合した質問ノードには意味制約として実体ノードの主辞の概念 ID を与える。従来では、図 4 に示すような「S80 はどのメーカーの車ですか。」と質問ノード（どの）と実体ノード（メーカー）が別の文節として分離している場合、回答ノード（ボルボ）と実体ノードが概念の類似性から照合してしまっていた。しかし、質問ノードと実体ノードを融合して質問ノード（どのメーカーの）とすることにより、質問ノードが回答ノードに正しく照合するようになり、回答の抽出精度を高めることができる。

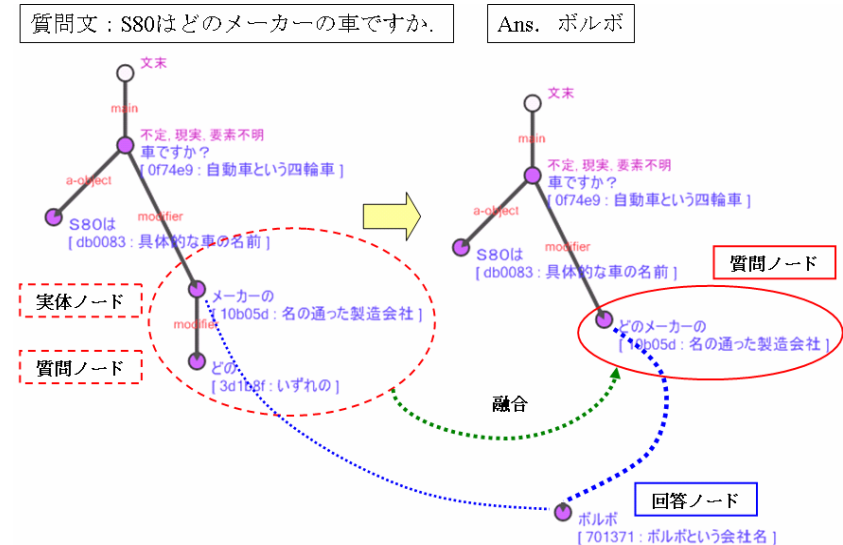


図 4 冗長型質問における質問ノードと実体ノードの融合

5. 知識文検索における改良

知識文検索における改良を以下に述べる。

5.1 知識文データベースのインデキシング

新聞記事より検索を行う場合は検索エンジンとして Lucene[16]を用いている。Lucene によるインデクス作成時にはあらかじめ新聞記事全てを意味解析し、その結果の意味グラフを知識文データベースとしてインデクスを作成する。インデクス・キーワードは意味解析時の形態素を単位として作成する。このため、質問文を意味解析した時と同じルールによって意味解析された結果を使用することにより語の区切りが統一される。インデクスを作成する際には以下のルールによってインデクス対象の語が選ばれる。

1. 文節ノードがリファレント（固有表現）を持っていた場合、リファレントを登録する。
2. 文節ノードの表記より助詞を取り除いた文字列を登録する。
3. 文節ノードを構成する形態素を登録する。ただし形態素に基本形があるものは基本形も登録する。尚ここで、その形態素が記号、助詞系列、括弧、接続詞、副詞、連体詞、判定詞、動詞の語尾にある“する”、“ある”、“なる”、“される”の場合は、それらが単独では意味をなさないので登録しない。

特に上記の3では語の基本形を利用することで“静かだ”と“静かな”といった活用の違いを吸収した検索を行うことができる。

インデクスの作成においては図5に示すように、語と深層格（語の役割）をペアにしたインデクスも登録する。深層格を含めることで「1979年に、米中が国交を正常化した。」という知識文は“正常化”という述語節に対して“time”，“agent”，“object”格で表される知識を持つことがわかる。これにより、答えを含まない知識文検索を阻止することができる。図6に示すような「米中が国交を正常化したのはいつ。」という質問文に対して、キーワードを「米中 国交 正常化」として知識文を検索しても、得られた知識文に国交が正常化したのが“いつ”なのかという答えが含まれているとは限らない。そこで、疑問詞“いつ”の疑問詞タイプが“時”であれば深層格 time を含めたキーワード“正常化+time”を指定することで、“正常化”が time 格を持つ知識文のみを検索できるようになる。言い換えれば、後続のグラフ照合による質問グラフに対する全体的な照合の前に、語の深層的な役割に基づく部分的な意味検索をインデクスを用いて高速に行うことができ、有効なスクリーニングになる。

知識文：1979年に、米中が国交を正常化した。

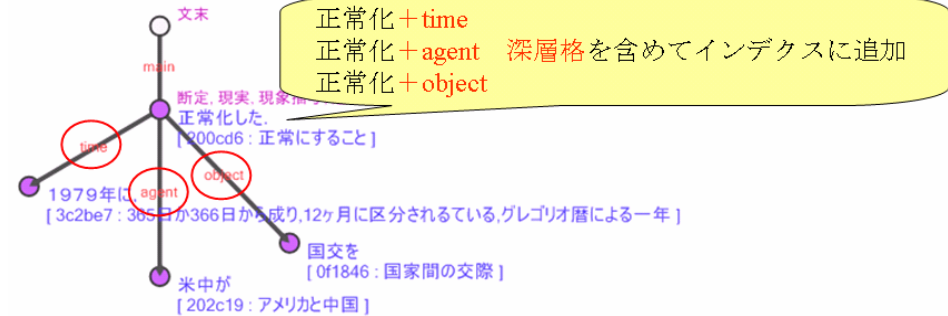


図5 語と深層格をペアにしたインデクス作成

質問文：米中が国交を正常化したのはいつ。

R：米中 M：国交 M：正常化

キーワード抽出

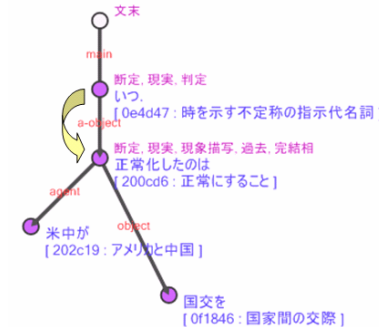


図6 深層格を含めたキーワード

5.2 質問文からのキーワード抽出

2006年以前[12,17,18]は、キーワードの組み合わせ作成時にはノード文節をキーワードの最小単位として扱っていた。しかし、検索エンジンで検索する際の検索文字列では文節単位では長すぎるという問題がある。例えば、「中国国家主席」というような複合語が一文節となった場合、「中国の国家主席」というような表現がデータベース上にあった場合に検索することができない。2007年度のシステム Metis2007[19]ではこの問題を解決するために、作成したキーワード組み合わせより検索文字列を作成する際には“中国”と“国家主席”というように形態素単位で分割していた。しかし、この方

法では形態素数が多くなると本来の複合語の意味から離れた記事も検索してしまう。そこで、リファレントを持つ複合語の場合にはまずリファレント全体をキーワードにする。その後形態素毎に品詞を確認することで、人名等の固有部分と役職・組織種別等の付属部分を分け、固有部分のみのキーワードと固有部分と付属部分の複合語からなるキーワードを OR 展開することで必要最小限の知識文を検索するキーワード文字列を作成することが出来る。

5.3 同意語展開

「飛行機を発明したのは誰ですか。」という質問において、キーワードは“飛行機”と“発明”の2つが抽出される。この2つのキーワードで知識文を検索するが、「航空機を発明したライト兄弟」という知識があったとしても“飛行機”ではなく“航空機”という表現になっているため検索で得ることができない。グラフ照合において“飛行機”と“航空機”は同じ概念であるためこの知識から回答を抽出することは可能である。このような知識も検索で得るために同意語展開を行う。同意語展開は、キーワードに対して EDR 辞書において同じ概念である語を同意語として検索を行うことである。同意語展開はキーワードとして抽出されたものの中でリファレントを持たない一般の語に対して行う。例としては“飛行機”というキーワードに対して“航空機”や“エアプレーン”といったものが同意語展開の結果得られる。得られた同意語の OR 展開を行うことにより同じ意味を持つ多くの知識文を検索できる。

5.4 結合された形態素中のブレの考慮

Metis2007 では表記のブレ（バイオリンとヴァイオリンなど）を複数の形態素が結合した語では行っていなかったため正答を含む知識文を検索できない場合があった。そこで、検索文作成の際に形態素が結合した語でもカタカナ・数字のブレ（ヴァ⇔バ、ヴェ⇔ベ、三⇔3など）を OR 展開して検索用キーワードを作成した。カタカナ表記のブレを考慮する変換ペア文字列については表3に示す。

表3 カタカナ表記のブレを考慮する変換ペア文字列

変換ペア文字列	
イ	ウイ
バ	ヴァ
ボ	ヴォ
ブ	ヴ
ジ	ヂ
ズ	ヅ
コ	クオ
ウオ	ウオ

ビ	ヴィ	
チ	ツイ	
カ	ケ	
エ	ウエ	ウエ
ベ	ヴェ	ベエ

6. 回答抽出における文キーワード類似度の付与

回答抽出において、Metis ではグラフ類似度が低くてもその知識文を含む記事全体にキーワードが多く含まれる場合は正解が存在する確率が高いので、グラフ類似度にキーワードが含まれる量に応じて最大 30 をキーワード類似度として加算して総合類似度（式 11）とする。また、キーワードと抽出された回答を分析した結果、知識文記事内だけでなく質問文と照合する知識文にキーワードを多く含んでいることが、より関連性のある知識文からの回答を優先できることが分かった。そこで、知識文記事に対するキーワード類似度（最大 30）を記事キーワード類似度（式 8）と、知識文に対する文キーワード類似度（式 9）に分けることにした。

$$\text{記事キーワード類似度} = \frac{0.5 \times \text{知識文記事内の } M \text{ 総数} + \text{知識文記事内の } R \text{ 総数}}{\text{質問文内のキーワード } (R, M) \text{ 総数}} \times 9 \quad (8)$$

$$\text{文キーワード類似度} = \frac{0.5 \times \text{知識文内の } M \text{ 総数} + \text{知識文内の } R \text{ 総数}}{\text{質問文内のキーワード } (R, M) \text{ 総数}} \times 21 \quad (9)$$

M : Must キーワード R : Referent キーワード

このとき、 R と M の重みは 2:1 とし、知識文記事内と知識文内の M 総数には 0.5 をかける。記事キーワード類似度と文キーワード類似度の重みは 30 を基に 3:7 に分けて、9 と 21 とする。結局、グラフ類似度に文キーワード類似度を足したものを文類似度（式 10）とし、それに記事キーワード類似度を足したものを最終的な総合類似度（式 11）として回答をこの総合類似度の順に提示する。これにより、キーワードを多く含む知識文の回答を優先して抽出することができる。

$$\text{文類似度} = \text{グラフ類似度} + \text{文キーワード類似度} \quad (10)$$

$$\text{総合類似度} = \text{文類似度} + \text{記事キーワード類似度} \quad (11)$$

7. 評価実験

NTCIR-6 CLQA (factoid 型 200 問) のデータに基づき毎日新聞記事 1998 年と 1999 年の 2 年分を知識源として実験を行った結果を表 4 に示す。

表 4 NTCIR-6 CLQA 実験結果

	検索	正解		正解抽出精度		回答抽出 精度
		1 位	5 位まで	1 位	5 位まで	
Metis2006	158/200	33	65	20.9% (33/158)	41.1% (65/158)	32.5% (65/200)
Metis2007	184/200	78	142	42.4% (78/184)	77.2% (142/184)	71.0% (142/200)
Metis2008	196/200	90	158	45.0% (90/196)	80.0% (158/196)	79.0% (158/200)

知識文の検索において、回答を含む知識文を 200 問中 196 問検索することができた。これは特に、知識文のデータベースにおけるリファレントを重視したインデキシングと、キーワード抽出において役割や組織名を含めたキーワード文字列作成が回答を含む知識文の取得に大きな効果を挙げたためだと考えられる。正解抽出精度においては、1 位で回答できたものが 45% (90/196)、5 位までに回答したものが 80% (158/196) であった。また、回答抽出精度は 79% (158/200) であった。これは、冗長型の質問の場合質問ノードと実体ノードを融合したことで回答ノードと照合するようになったことが特に有効であったためである。

8. 結論

本研究では、質問文解析フェーズと検索フェーズを中心に従来のシステムを改良した。評価実験の結果、知識文検索および正解抽出精度ともに従来のシステムを上回るものとなった。プロトタイプであった Metis2006 と比較すると、Metis2008 では回答抽出精度が 79% と当時の精度を 2 倍以上に向上したことになる。また、この精度は NTCIR-6 CLQA のコンテスト中最高位の 59.5% を上回るものだった。

参考文献

- [1] Boris Katz: Using English for Indexing and Retrieving, MIT Press, Vol.2 (1990).
[2] Julian Kupiec: MURAX: A robust linguistic approach for question answering using an on-line encyclopedia, In Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp.181-190 (1993).

- [3] 情報処理学会研究報告, 2003-NL-158, pp.69-74 (2003).
[4] 佐々木裕, 磯崎秀樹, 鈴木潤, 国領弘治, 平尾努, 賀沢秀人, 前田英作: SVM を用いた学習型質問応答システム SAIQA-II, 情報処理学会論文誌, Vol.45, No.2, pp.635-646 (2004).
[5] 清水 慧, 秋葉 友良, 藤井 敦: 統計翻訳に基づくパッセージ検索の言語横断質問応答への適用, 言語処理学会第 13 回年次大会発表論文集, E5-5, pp.1176-1179 (2007).
[6] CLQA: <http://clqa.jpn.org/>
[7] 尾見孝一郎, 原田実, 岩田隆志, 水野高宏: 日本語文章からの意味フレーム自動生成システム SAGE(Semantic frame Automatic GEnerator)の開発研究, 人工知能学会第 13 回全国大会論文集, pp.213-216 (1999).
[8] 原田実, 水野 高宏: EDR を用いた日本語意味解析システム SAGE, 人工知能学会論文誌, Vol.16, No.1, pp.85-93 (2001).
[9] 原田実, 田淵和幸, 大野博之: 日本語意味解析システム SAGE の高速化・高精度化とコーパスによる精度評価, 情報処理学会論文誌, Vol.43, No.9, pp.2894-2902 (2002).
[10] 杉村和徳, 山本哲哉, 木村健太郎, 鳥居隼, 韓東力, 原田実: 意味解析システム SAGE の精度向上と利便性の向上, 情報処理学会第 67 回全国大会論文集, 1J-02, pp.67-68 (2005).
[11] 梅澤俊之, 西尾華織, 松田源立, 原田実: 意味解析システム SAGE の精度向上とモダリティの付与と辞書更新支援系の開発, 言語処理学会第 14 回年次大会発表論文集, pp.548-551 (2008).
[12] 竹原一彰, 安部建助, 安田智成, 韓東力, 原田実: 質問応答のための質問文と知識文の間の意味ベースでの精密な照合方式, 情報処理学会第 66 回全国大会論文集, 6U-03, pp.173-174 (2004).
[13] Minoru Harada, Yuhei Kato, Kazuaki Takehara, Masatsuna Kawamata, Kazunori Sugimura, and Junichi Kawaguchi: QA System Metis Based on Semantic Graph Matching, Proc. of the 6th International Conference on NII Test Collection for IR Systems(NTCIR6), Tokyo, Japan, pp.448-459 (2007).
[14] QAC: <http://www.nlp.is.ritsumei.ac.jp/qac/>
[15] EDR: <http://www2.nict.go.jp/r/r312/EDR/>
[16] Lucene: <http://incubator.apache.org/lucene.net/>
[17] 加藤裕平, 古川勇人, 蒲生健輝, 韓東力, 原田実: WEB 検索による知識文の獲得と意味グラフ照合推論による質問応答システム Metis, 情報処理学会第 67 回全国大会論文集, 1G-06, pp.11-12 (2005).
[18] Dongli HAN, Yuhei KATO, Kazuaki TAKEHARA, Tetsuya YAMAMOTO, Kazunori SUGIMURA, and Minoru HARADA: QA System metis based on web searching and semantic graph matching, Proc. of the 4th International Conference on Intelligent Information Processing, Adelaide, Australia, pp.123-133 (2006).
[19] 久保田裕章, 平塚飛将, 吉川ひかる, 松田源立, 原田実: 質問応答システム Metis の回答精度向上 - 検索フェーズの改良を中心として -, 言語処理学会第 14 回年次大会発表論文集, A5-5, pp.1017-1020 (2008).