

インターネットへの高精度な質問応答システムの開発

代表研究者 原 田 実 青山学院大学理工学部情報テクノロジー学科教授
 共同研究者 松 田 源 立 青山学院大学理工学部情報テクノロジー学科助教

1 はじめに

近年、自然文で書かれた質問文に対して膨大な量の文章より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おり、NTCIRのQuestion Answering Challenge (QAC) [8]やCross Language Question Answering (CLQA) [9]等の評価型ワークショップなどで研究成果が公表されている。

多くの手法が発表されているが、それらの基本的考え方は、質問文に対し、その内容的な類似性から答えを持ちそうな知識文をインターネットや新聞から検索し、疑問詞に対応する箇所を知識文から取り出し回答として返すものである。

初期の研究では、質問文と知識文の類似性判定はTF/IDF法で行われていたため、回答精度は非常に悪かった。倉田[1]らの研究では、回答候補を抽出した後に係り受け解析によって得られた文節をノードとするグラフの構造を用いてノード間の距離を求め回答を抽出している。ノード間の距離を求める際に質問文から抜き出された検索キーワードとの距離で回答候補のスコアを求め、回答候補より回答を抽出している。ただ、距離計算において係り受け関係のみを考慮しており、ノード間の役割的关系を考慮していない。これにより知識文に余分な修飾表現がなされていた場合においてノード間距離が増加して回答順位が下がり回答を正しく抽出できない場合が発生する。

村田[2]らの研究では、質問文及びデータベース中より抽出した知識文を係り受け解析し、構文情報を利用して2つの文を照合し、文同士の類似度を計算して回答を抽出している。この場合照合した2文の文節間の対応をとり疑問詞と対応付けられた文節を回答として抽出している。しかし、この方法では構文情報のみで文節間の意味的关系については考慮されていない。また、文の意味解析を行っていないために対応する文節間の類似度計算が構文情報を中心に行われており、意味的な類似度を求められない。

以上のように文を形態素あるいは構文的な情報のみで回答抽出を行うと文の意味的内容を正確に把握することが出来ない。この結果、質問文中の語との意味的な類似性や語間の意味的な関係を考慮せずに知識文中の語との対応をとるため、誤回答を抽出してしまうことが多い。

2 研究目的

本研究では、自然文で与えられた質問文に対し意味解析を行い、質問文と知識文(インターネットや新聞中の文で回答を含む可能性のある文)との意味的対応を十分正確に照合しながら回答を抽出するシステムMetisを開発する。Metisでは質問文(例:「誰が何時ペスト菌を発見しましたか?」)と知識文(例:「北里柴三郎が1894年に香港でペスト菌を発見した、」)が表す内容を精密に照合するために、それぞれに対して従来の自然語処理である形態素解析(長尾・黒橋らのjuman[10]を使用)と係り受け解析(黒橋・河原らのknp[10]を使用)を行うだけでなく、原田研究室で開発したSage[4]によって意味解析と照応解析を行いその結果を図1上段に示すような意味グラフの形式で出力する。意味グラフでは、各語にEDR電子化辞書中の語意(6桁の16進数)が割り振られ、語間の意味(役割)関係は30数共通部分の質問グラフに対する相対的大きさとして測定する。例えば、先の2文間における類似共通部分グラフは、図1下段に示す、4つのノード対(例:発見した、←(0.92:0.92)→見つけましたか、)と3つのアーク対であり、ノード対の語意的類似度はEDRの概念体系木上での共通上位概念までの距離に基づいて計算し(例では0.92で、:の後の0.92はモダリティの違いによる低減後の値、事例ではモダリティが同じなので同値となっている)、アーク対の類似度は我々が定めた深層格のどの類似グループに入っているかによって計算する。これらの類似度を合計し、質問グラフのノード数やアーク数で割ってノードグラフ類似度とアークグラフ類似度を求め、その平均をグラフ類似度とする。これによって、2つの文における、「何が何処で何を何時どんな風にどうした」のあらゆる観点からそれぞれの対応する語の類似性を評価するという、人における文の類似判定とほぼ同様の判定を

行うことができる、

3 質問応答の流れ

図2のシステムフローで示すように、Metisではまず意味解析をして得た質問グラフより抽出した検索キーワードを用いて新聞データやWebを検索し知識文を獲得し、これを同様に意味解析して知識グラフとし、これらと質問グラフを照合して回答を抽出する。

3.1 質問タイプ分類

Metisでは人名、数量といったものを問うfactoid型の質問だけでなく、why, how, definitionといった質問を含む全ての質問に対して回答するため、処理のはじめに入力された質問を質問タイプ別に分類する。本システムでは表1に示すようにfactoid型の質問を12タイプに分類し、さらにwhy, how, definition型の質問をそれぞれ1タイプに分類する。

3.2 検索キーワードの抽出

検索キーワードは質問グラフのノード(文節)単位で抽出を行う。質問グラフより検索キーワードを抽出する際には、NormalキーワードとMustキーワードの2タイプに分けて抽出する。Mustキーワードとは知識ベースの検索時に必ず指定する質問の核となるキーワードで、質問グラフから1つ以上選出する。例えば「奈良県にある「東大寺南大門」の両脇に立つ像は何か、」という質問文の場合、Mustキーワードとして「奈良県」、「東大寺南大門」、「像」の3つが抽出

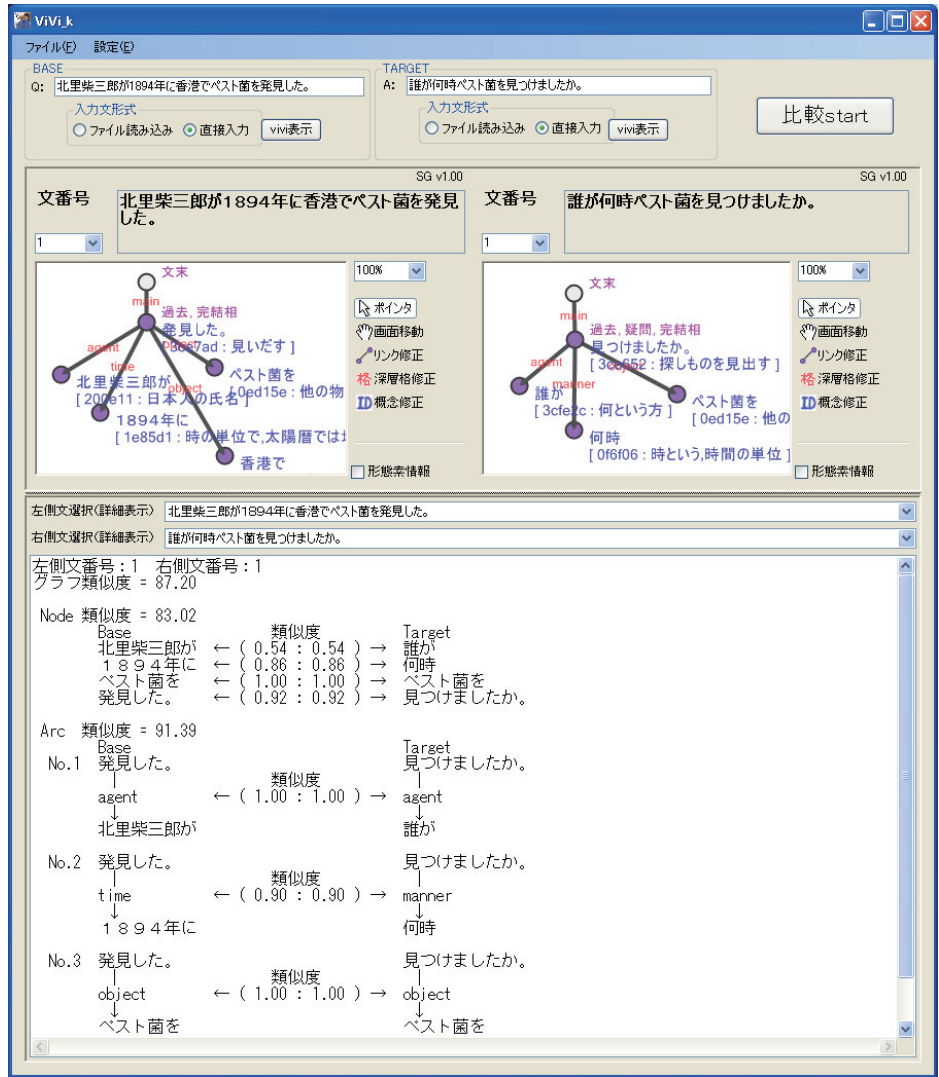


図1 質問文と知識文との意味的対応

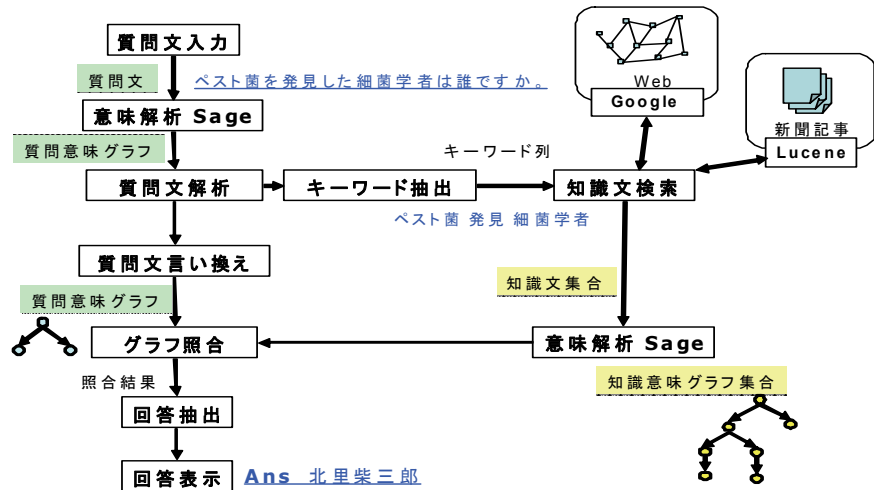


図2 METISにおける処理の流れ

され、Normal キーワードとして「両脇」、「立つ」、
「像」の3つが抽出される。このとき、「両脇」、
「立つ」、像」という Normal キーワードのみで検索
を行うと、求める知識とは関係ない不要な知識が多
く検索されてしまうことがあるので、知識ベース検
索の際には Must キーワードを含むキーワード群で
検索を行うことにより、不要な検索を減らすことが
できる。

3.3 データベース検索による知識文の取得

本システムでは知識のデータベースを Web と新聞
記事の2つより選択して検索を行う。Web 検索の際
には検索エンジン Google[6]を用いる。新聞記事よ
り検索を行う場合は検索エンジン Lucene を用いる。
Lucene によるインデックス作成時にはあらかじめ新聞
記事全てを意味解析し、それをデータベースとしてイン
デックスを作成する。インデックス・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。データ
ベースより検索を行う際のキーワードはグラフの文節
ノードを構成する形態素単位で抽出する。これによ
り、「中国国家主席」というような複合語が文節ノ
ードとなった場合、「中国」と「国家主席」が検索
キーワードになり、「中国の国家主席」というよう
な表現も検索できるようになる。また、西暦を表
記する場合において、「1982年」を省略して「82
年」と表記している場合が多々ある。そこで、西
暦がキーワードとなっている場合、「1982年」と
「82年」の2通りの検索を行うこととする。

また、本システムではグラフ照合の際に、ノ
ードの語意については概念体系木を用いて類似度
を求めているため、概念が同じならば「発見」と
「みつけ出す」というように表記が違っても同
じものであると判断することができる。しかし、知
識文をデータベースより収集する時点では「発見
」というキーワードで検索を行うため「みつけ出
す」という文は見つけることができない。そこで
本システムでは EDR を利用し同じ語意のものを
キーワードの言い換え語として使用し、検索時に
「発見」と「みつけ出す」の2通りでの検索を
行っている。このようにして同意語をキーワード
に含めることで、より多くの正解を含んでいそ
うな知識文を収集することができる。

また、本システムではグラフ照合の際に、ノ
ードの語意については概念体系木を用いて類似度
を求めているため、概念が同じならば「発見」と
「みつけ出す」というように表記が違っても同
じものであると判断することができる。しかし、知
識文をデータベースより収集する時点では「発見
」というキーワードで検索を行うため「みつけ出
す」という文は見つけることができない。そこで
本システムでは EDR を利用し同じ語意のものを
キーワードの言い換え語として使用し、検索時に
「発見」と「みつけ出す」の2通りでの検索を
行っている。このようにして同意語をキーワード
に含めることで、より多くの正解を含んでいそ
うな知識文を収集することができる。

3.4 正答を含まない知識文検索の排除

従来の Metis では、質問文から抽出したキー
ワードのみを用いて知識文を検索していたが、そ
れだけでは回答のない無駄な知識文を検索して
しまう可能性がある。例えば、「米中が国交を正
常化したのはいつですか?」という質問文に対
して、検索文を「米中 国交 正常化」として
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。このような表層的な検索ではなく深
層的な検索を行うために、深層格（語の役割）
を含めた検索を行う、

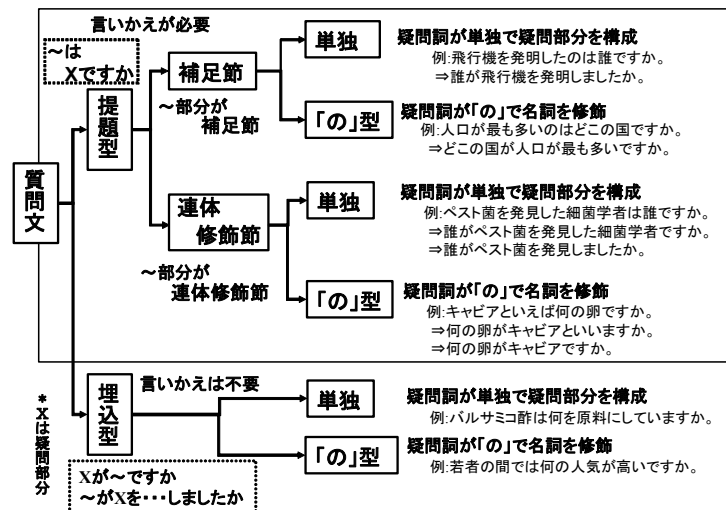
(1) 語と深層格をペアにした Index 作成

語と深層格をペアにした Index の作成を行
うことで、正答を含まない知識文検索を阻止す
る。Index 対象となる語は前節で述べた通り
である、それらの語に、深層格を加えて登録す
る。ルールは以下の通りである、
R1) ノードの出力辺の深層格を、登録する語
の後ろに「+深層格」として登録する、
R2) ノードの入力辺の深層格を、登録する語
の前に「深層格+」として登録する、
例えば、図4に示すように「1979年に、米中
が国交を正常化した、」という知識グラフに対
しては、通常

表1 質問分類

質問タイプ	質問例
誰	電池を発明した、イタリアの物理学者は誰ですか。
何時	「マッチ売りの少女」はいつの日の話ですか。
どこ	インドネシアの首都はどこですか。
どこ	国際デジタル通信はこの会社に買収されましたか。
何X	アメリカ合衆国で最も面積の大きい州は何州ですか。
何の	「ボジョレー・ヌーボー」とは何のお酒ですか。
どんな	「キューブ」といえばどんな形ですか。
どの	どのチームを吸収合併しましたか。
どのくらい	富士山の高さはどのくらいですか。
どのように	マカオはポルトガル語でどのように表しますか。
いくら	1998年の日本の経常黒字はいくらですか。
どれくらい	どれくらいの人が携帯電話を持っていますか。
why	なぜ傘を持っていくのですか。
how	どうやってアメリカに行きましたか。
definition	K-1とは何ですか。

図3 質問文の言い換え対象の分類



のインデックスに加えて、正常化+ti, 正常化+ag, 正常化+ob, ti+1979年, ag+米中, ob+国交の6つが登録される, このように深層格込みのインデックスを登録することにより, この文が「正常化」の「時」「主体」「相手」を知識に持つことがわかる,

(2) 深層格を含めたキーワード検索

深層格を含めたキーワード検索を行うことで, 質問タイプに対応した回答を含む知識文のみを検索できるようになり, 高速な意味検索が可能となる,

まず, 入力された質問文に対して質問文解析とキーワード抽出を行い, 質問文解析において質問タイプが決定される, そこで, 質問グラフ内の疑問詞に係る述語節ノードに図5に示すように質問タイプに対応した深層格を追加してこれキーワードにして検索文を作成する,

3.5 質問文の言い換え

本システムでのグラフ照合の際, 主述語としては「～したのは…ですか」(提題型)というような表現ではなく「…が～したか」

(埋め込み型)のような表現にした方が知識文との類似度が高まる場合がある. 例えば, 「ペスト菌を発見した細菌学者は誰ですか」という質問に対し, 「北里柴三郎がペスト菌を発見した」という文は係り受けに関するグラフ構造の違いから文同士の類似度が低くなる. しかし, 「誰がペスト菌を発見したか」というような表現への言い換えを行えば文の類似度が高くなる. このようなグラフ構造の違いによる類似度の低下を防ぐために質問文に対して言い換え処理を行う. 質問文の形式を大きく提題型と埋め込み型に分類し, 提題型と分類された質問文に対して言い換え処理を行う. 提題型は図3に示すようにさらに補足節型と連体修飾節型の2つに分類し, それぞれに対応した言い換え処理を行う. ・補足節型の言い換えについては, 例として「飛行機を発明したのは誰ですか」という質問文を「誰が飛行機を発明しましたか」という文に意味グラフ上での変形を行う.

・連体修飾節型の言い換えについては, 例として「ペスト菌を発見した細菌学者は誰ですか」という質問文を「誰がペスト菌を発見しましたか」や「誰がペスト菌を発見した細菌学者ですか」という文に意味グラフ上での変形を行う.

このように変形した質問グラフと変形前の質問グラフの両方を用いて知識グラフとの照合を行い知識グラフとの類似度を調査する.

3.6 質問グラフと知識グラフの照合

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間の概念類似度を計算し, 概念類似度が閾値を超えた質問グラフノードとそれに対応する知識グラフノード間の概念類似度(これをノード類似度という)を合計してノードグラフ類似度とする. 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合計しアークグラフ類似度とする. 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とする. 2つ

知識文: 1979年に、米中が国交を正常化した。

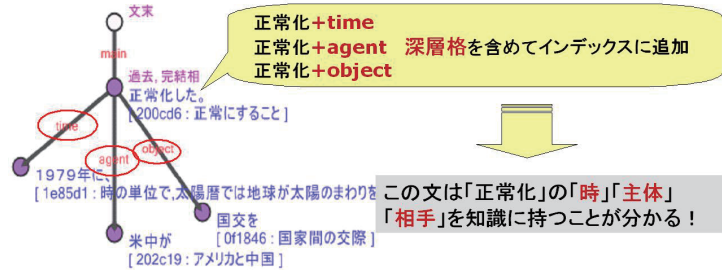


図4 語と深層格をペアにした Index 作成

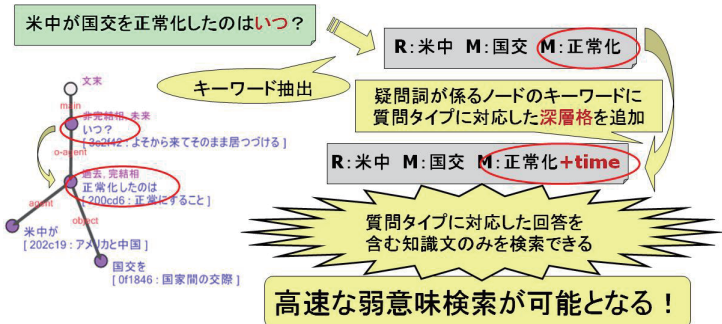


図5 深層格を含めたキーワード検索

表2 深層格グループ

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

のノードの類似度はそれぞれのノードが持つ概念の類似度とする（固有名詞を表す場合はその表記による類似度を基に計算する）。2つの概念C1とC2の概念類似度はEDRの概念体系木上での共通上位概念c(c1, c2)までの距離に基づいて以下の式によって求める。

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

与えられた2辺のアーチ類似度はそれぞれのアーチが持つ深層格が表2に示す我々が定めた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。ノード及びアーチの類似度から以下の式でグラフ類似度を計算する。

$$\text{グラフ類似度} = \text{ノードグラフ類似度} + \text{アーチグラフ類似度}$$

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

$$\text{アーチグラフ類似度} = \frac{\sum(\text{アーチ類似度} \times \text{ムード得点})}{\text{質問グラフのアーチ数}} \times 50$$

ここで、モダリティ得点とはノードが持つ「断定」や「疑問」や「過去」といったモダリティを比較して決められる得点である。例えば、「発見した」と「発見していない」や「発見したい」といったノードを比較しようとした場合、語意は全て「発見」という概念であるため概念類似度は高い。しかし「発見した」と「発見していない」ではまったく逆のことを言っている。このような場合に1以下のモダリティ得点を掛けることによりノード類似度を低くし、人間の感性にあった類似度を与えることができる。

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3に示すような代替概念を用いる。例えば、質問タイプが「誰」であるならば「人名」といった質問内容に則した代替概念を用いる。こうすることにより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。

表3 質問タイプ別代替概念

質問タイプ	与えられる代替概念
誰	人の名称, 人の呼称, 人間
何時	時, 時間点, 数量, 計量の単位
どこ	組織, 地名, 国家
どこの	質問箇所の上位ノードの概念
何X	質問箇所の主辞, 副主辞の概念
何の	具体物, 抽象物, 場所, 自立活動体, 状態
どんな	質問箇所の上位ノードの概念
どの	質問箇所の上位ノードの概念
どのくらい	数量, 計量の単位, 状態
どのように	事象, 物事, 抽象物, 状態
いくら	数量, 計量の単位
どれくらい	数量, 計量の単位, 状態

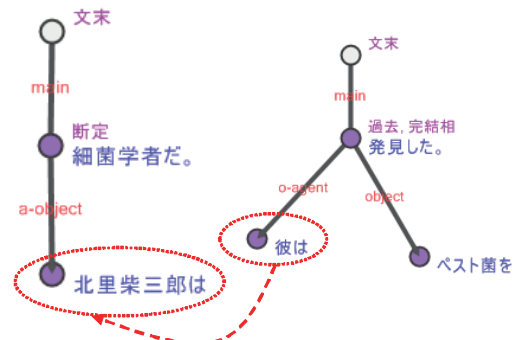


図6 指示代名詞

3.7 照応関係の利用による複数文での照合

本システムで利用している意味解析システムSageでは文の照応関係についても解析することができる。照応関係を利用することにより、本システムでは複数文からの知識による照合を行っている。

Sageが解析する照応関係には指示代名詞及びゼロ代名詞がある。図6に示すように指示代名詞では「北里柴三郎は細菌学者だ。彼はペスト菌を発見した。」という文において「彼」が「北里柴三郎」であるという情報が付与される。一方、図7に示すように、ゼロ代名詞では「北里柴三郎は細菌学者だ。ペスト菌を発見した。」という文において「ペスト菌を発見した」という文に対して「北里柴三郎」がその動作主体であるという情報が補完される。これらの情報をグラフ照合の際に使用することにより複数文からの回答抽出を可能としている。

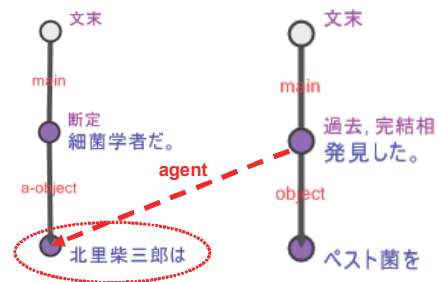


図7 ゼロ代名詞

3.8 回答抽出

グラフ照合の結果、質問グラフとの類似度が高い知識グラフより回答を抽出する。回答抽出の基本方針は

factoid型の場合は質問箇所ノードと照合した知識ノードを回答として抽出し、why型とhow型とdefinition型では質問グラフ内の質問の主題を表す主題ノードを決定し、主題ノードと照合した知識ノード(主知識ノード)より回答として妥当な深層格で結合する知識ノード(根拠ノード)を決定し、この根拠ノードを根とした部分木を回答として抽出する。回答抽出方法の詳細は次節で示す。

(1) factoid型の回答抽出

factoid型の質問の場合、グラフ照合の際に述べたように「誰」や「どこ」といった質問ノードと照合した知識ノードを回答として抽出する。回答抽出の際には回答として抽出するノードを修飾している(modifier格で結合している)ノードが存在した場合はそれも含めて抽出する。例えば図8に示すように、「東海村ウラン処理施設事故で、大内久さんは一般的な年間被ばく許容量の何倍の放射線を浴びたのですか?」という質問に対して「国内の原子力施設で初めて起きた臨界被ばく事故で、被ばく量は、一般の人が1年間に浴びる放射線の許容限度の1万8000倍に当たる約18シーベルトと推定され、国内で過去最大級だった。」という知識が存在した場合に回答として「1万8000倍」を抽出する。このとき文中で「許容限度の」が「1万8000倍」を修飾しているので回答を「許容限度の1万8000倍」として抽出する。

(2) why型の回答抽出

why型の質問の場合、照合した知識グラフより質問グラフ中の(検索キーワードとなっている)主述語ノードを主題ノードとし、これと照合する主知識ノードと「reason」、「cause」、「manner」、「sequence」、「location」、「sequence」という理由を表現する深層格で結ばれている根拠ノードを根とする知識グラフ内の部分木の集合を抽出する。例としては図9に示すように「なぜインドとパキстанは対立しているのか。」という質問文に対して「パキстанはタリバン政権を承認する唯一の国だが、インドは反タリバン連合(北部同盟)に近く、

アフガン情勢をめぐってインドとパキстанは微妙に対立している。」という知識文から、主知識ノードである「対立している」より「sequence」格、「cause」格で結ばれている根拠ノードを根とする部分木を取り出して「パキстанはタリバン政権を承認する唯一の国だが、インドは反タリバン連合(北部同盟)に近く、アフガン情勢をめぐって」が回答として抽出される。また、単一の知識グラフ内に理由を表す深層格が無い場合においても、Sageの文脈解析により後続の文との関係が「理由」である場合は後続の文を回答として抽出する。この場合は、知識文が「インドとパキстанは対立している。

なぜならパキстанはタリバン政権を承認する唯一の国だが、インドは反タリバン連合(北部同盟)に近く、アフガン情勢をめぐっているからだ。」という場合に、「インドとパキстанは対立している」が文間深層格「理由」で後続の文の「なぜならパキстанはタリバン政権を承認する唯一の国だが、インドは反タリバン連合(北部同盟)に近く、アフガン情勢をめぐっているからだ。」と接合しているためこれを回答として抽出する。

(3) how型の回答抽出

how型の質問の場合、質問グラフの主述語ノードを主題ノードとし、主知識ノードと「implement」、「sequence」、

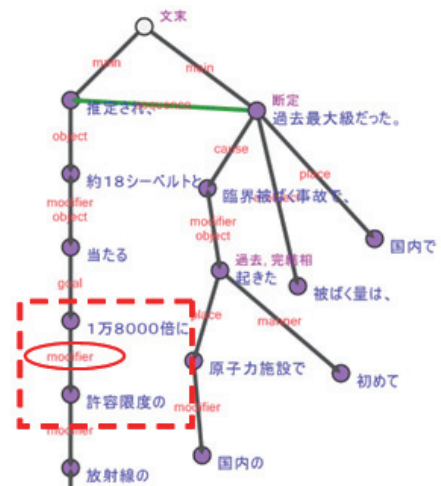


図8 factoid型回答抽出

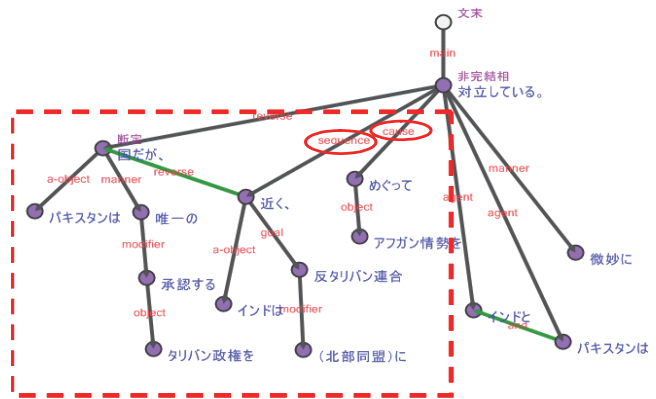


図9 why型回答抽出

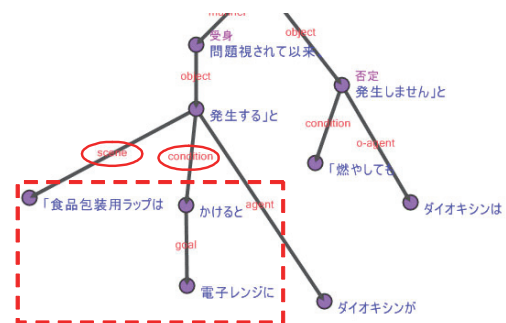


図10 how型回答抽出

「condition」, 「manner」, 「scene」という手段や方法を表現する深層格で結ばれている根拠ノードを根とする知識グラフ内の部分木の集合を抽出する. 例として図 10 に示すように, 「ダイオキシンはどのようにして発生しますか. 」という質問に対して「実際, 「食品包装用ラップは電子レンジにかけるとダイオキシンが発生する」と問題視されて以来, …」という知識が存在した場合は主知識ノードの「発生すると」の「condition」格と「scene」格で結ばれている根拠ノードを根とする部分木からなる「食品包装用ラップは電子レンジにかけると発生する」が回答として抽出される.

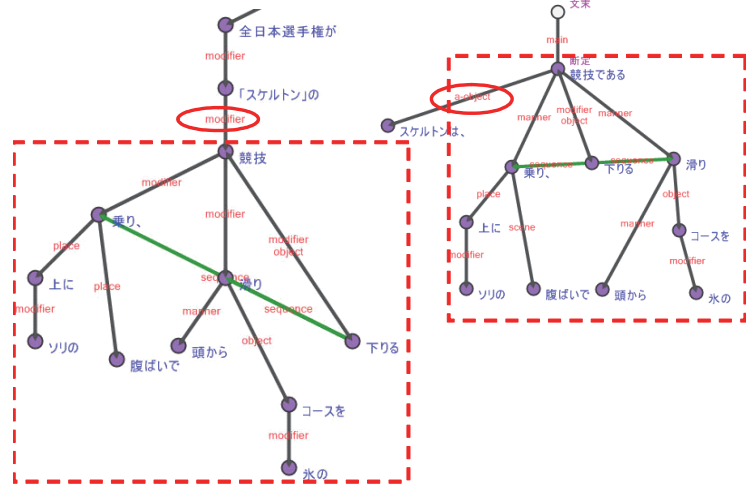


図 11 definition 型回答抽出

(4) definition 型の回答抽出

definition 型の質問の場合, 定義を問われている語を主題ノードとし, 主知識ノードと「modifier」格で結ばれている根拠ノードを根とする知識グラフ内の部分木を抽出する. 例としては図 11 に示すように, 「スケルトンとはどのような競技ですか. 」という「スケルトン」という語を問う場合において「ソリの上に腹ばいで乗り, 頭から氷のコースを滑り下りる競技「スケルトン」の全日本選手権が 10 日, …」という文があった場合, 「スケルトン」の「modifier」格を根とする部分木である「ソリの上に腹ばいで乗り, 頭から氷のコースを滑り下りる競技」が回答として抽出されることになる. また, 主知識ノードが「a-object」で係っている語があった場合はその語を根拠ノードとし, 部分木を抽出する. この場合は「スケルトンはソリの上に腹ばいで乗り, 頭から氷のコースを滑り下りる競技である」という知識に対して「ソリの上に腹ばいで乗り, 頭から氷のコースを滑り下りる競技」が回答として抽出されることになる.

4 評価実験

4.1 クイズミリオネアに対する評価実験

クイズミリオネアの質問文 100 文に対してインターネットからの知識検索による質問応答を行った結果を表 4 の第 1 行目に示す, 有効な知識文の検索率 (再現率) が 96%, 回答精度 74%と良い精度を示した, 質問文や知識源が違うので厳密には比較にはならないが, 現在提案されている質問応答システムの多くは正解率が 50%台であるので, 非常に良い結果といえる, これは, 意味グラフ照合結果からの回答抽出という基本的な方式に加え, 意味制約ルールの詳細化や回答抽出の精度向上, グラフ照合における数詞類似度の追加など, 様々なフェーズにおいて改良を加えた結果, 達成したものといえる, これ以上の検索率向上を得るには, 動詞の言い換え処理(「暗殺された」を「死亡した」に言い換える)等が必要である,

4.2 NTCIR CLQA による評価実験

factoid の質問が対象となっている NTCIR CLQA のテストコレクションによる実験結果を表 4 の 2 行目に記す, 知識源は 1998 年と 1999 年の毎日新聞記事である, テストコレクション 200 問を対象として行った結果, 知識文の検索率 (再現率) が 92%, 正解抽出率が 77.2%という精度を示した, この精度は 2007 年 NTCIR の CLQA コンテスト中最高位の 59.5%を上回るものだった, ここで, 正解抽出精度がクイズミリオネアよりも低いのは 1 つの知識の文章量の違いよるところが大きいと

表 4 METIS の精度

	検索	正解		正解抽出精度		回答精度	
		1 位	5 位まで	1 位	5 位まで	1 位	5 位まで
クイズミリオネア	96/100	49	74	51%(49/96)	77%(74/96)	49%	74%
NTCIR-CLQA	184/200	78	142	42.4%(78/184)	77.2%(142/184)	39%	71%
NTCIR-QAC	81/100	21	42	25.9%(21/81)	51.9%(42/81)	21%	42%

考えられる，Web を使った検索の場合，利用する知識は Web ページの要約文を使用しているため文章が非常に短く，文照合が容易に行うことができる，しかし，新聞記事を知識源とした場合，1つの知識（新聞記事）の文章量が多いため係り受け等の意味解析の精度も短い文に比べ低くなる，このため，グラフ照合の結果を重視している本手法では意味解析の精度の低下の影響が大きく精度を下げる結果となっていると考えられる，

4.3 NTCIR QACによる評価実験

理由と方法と定義を問う質問が対象となっている NTCIR QAC のテストコレクションによる実験結果を表 4 の 3 行目に記す，テストコレクション 100 問を対象として行った結果，知識文の検索率（再現率）が 81%，正解率 42% という精度を示した，ここでも知識文の検索率は高く，検索方法の改善の効果が表れているといえる，正解抽出精度の 52% という値も factoid の質問に比べて低いものの，回答が文章という抽出が難しい対象であることを考えると高い精度であるといえる，

5 まとめ

評価実験において，正解になったものの中で第一位の回答が占める割合が高い．これは質問文と知識文を意味解析し両者を意味グラフとして照合し，最も類似度の高い知識文から回答部分を抽出するという，人が回答を得ると同様の方法で回答を抽出した結果であるといえる．また，この基本的手法が効果的に機能するために，質問文が何を回答として求めているのかを解析するための意味制約ルールの詳細化，知識文から回答部分を抽出する機構の詳細化，グラフ照合における数詞類似度の追加など，様々なフェーズにおいて改良を加えた，さらに，知識文検索においては，深層格込み検索を導入し回答が含まれない知識文の順位を下げたことで，答えを含む知識文を上位に検索できるようになった，ただし，精密な照合を行っているため意味解析の精度に大きな影響を受けてしまい，長文での照合では係り受け解析を間違えることがあり精度を下げる原因となってしまっている．また，本研究では Factoid 型の質問文に対する回答精度の向上を重視したが，今後は NTCIR の QAC といった，理由と方法と定義を回答として要求する質問に対する回答抽出の精度向上も目指すべきである，

【参考文献】

- [1] 倉田岳人，岡崎直観，石塚満：係り受け関係に基づくグラフ構造を用いた質問応答システム，情報処理学会自然言語研究会報告NL-158，pp. 69-74 (2003)．
- [2] 村田真樹，内山将夫，伊佐原均：類似度に基づく推論を用いた質問応答システム，情報処理学会自然言語研究会報告NL-135，pp161-188 (2000)．
- [3] 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，第2分冊pp. 173-174 (2004)．
- [4] 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，第2分冊，pp. 67-68 (2005. 3)．
- [5] Lucene.net, <http://incubator.apache.org/lucene.net/>.
- [6] Google Web APIs (beta), <http://www.google.com/apis/>.
- [7] EDR, <http://www2.nict.go.jp/r/r312/EDR/>.
- [8] Question Answering Challenge(QAC), <http://www.nlp.is.ritsumei.ac.jp/qac/>.
- [9] Cross Language Question Answering (CLQA) <http://clqa.jpn.org/>.
- [10] Juman, knp, <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/top.html>.
- [11] クイズミリオネア編：完全攻略!クイズ\$ミリオネア，株式会社フジテレビ出版(2002)．

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
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年5月
クレーム内容の自動分類	言語処理学会第14回年次大会発表論文集, pp. 293-294	2008年3月
要約システム ABISYS の精度向上とクレーム要約への対応	言語処理学会第14回年次大会発表論文集, pp. 492-495	2008年3月
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年3月
意味解析に基づく照応解析システム ANASYS の精度向上と大規模テキストコーパスによる評価実験	言語処理学会第14回年次大会発表論文集, pp. 552-555	2008年3月
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, pp. 1017-1020	2008年3月