

算数問題のような数量の論理関係を含む質問文の意味理解と解決
 Meaning Understanding and Solution of Question Sentence
 Including Logical Relation of Amount like Arithmetic Problem

阿部 一貴†
Kazutaka Abe

吉村 枝里子†
Eriko Yoshimura

渡部 広一†
Hirokazu Watabe

河岡 司†
Tsukasa Kawaoka

1. はじめに

コンピュータと人間との円滑な会話を実現するためには、人間のような常識に基づく会話の理解が必要である。会話には数量を含むものがあり、これの理解には言葉の意味の他、数量の論理関係の理解も必要となる。

本稿では、数量を含む文章の代表として、演算が必要な算数問題の意味理解を対象とする。その一例として速度問題について意味理解を行うシステムを構築する。辞書や概念ベースなどを用いた関連度計算方式、また速度問題に関しての知識を格納した知識ベースなどを用い、自然言語文章で記述された問題を算数問題として理解し解答を求める方式を考案する。

2. 関連技術

2.1 概念ベース

概念ベース^[1]とは語(概念)と意味(属性)のセットを約12万語蓄積したものである。この概念と属性のセットには、その重要性をあらわす重みが付与される。ある「概念A」は、その概念と関連が深く、その概念の意味となるその属性の重要性を表す「 w_i 」の組の集合で表される。概念Aは以下のように表される。

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_n, w_n)\}$$

また、属性 a_i を概念Aの一次属性と呼ぶ。一次属性 a_i を一つの概念と見なせば、 a_i からさらにその一次属性を導くことができ、 a_i の属性 a_{ij} を概念Aの二次属性という。これを展開していくと、一つの概念Aはn次属性まで持つことができる。

2.2 関連度計算

概念Aと概念Bの関係の深さを定量的に表すのが関連度計算方式^[2]という方法である。関連度は、0以上1以下の連続的な数で表され、概念同士の関連が大きいほど関連度は高くなる。この関連度を求める計算は、それぞれの概念を二次属性まで展開し、その重みを利用した計算によって最適な一次属性の組み合わせを求め、それらが一致する属性の重みを評価することで算出する。

3. 算数問題解決システム

算数問題は、定式化された問題(特殊問題と呼ぶ)と一般問題に分けられる。特殊問題とは、SPIの非言語能力問題のように公式を利用して解くことのできる数列問題、速度問題、確率問題、N進数問題や仕事算などである。また一般問題とは、小学校の算数問題や日常会話における数量を扱う会話など、一般的な四則演算などを利用して解く問

題である。例えば、「林檎が5個あります。蜜柑が10個あります。合わせていくつの果物がありますか?」というような問題である。

このような算数問題の言葉の意味、数量の論理関係から求められている解を導くことのできるシステムが「算数問題解決システム」^[3]として構築されている(図1)。

算数問題解決システムでは、特殊問題に関しては一部の問題にしか対応できていない。そこで、特殊問題の一つである速度問題を意味理解し、解決できるシステムを構築することによって、算数問題解決システムは幅広い分野で適用できると考えられる。

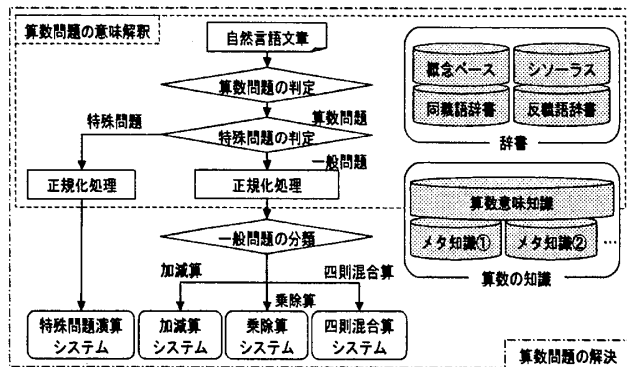


図1 算数問題解決システムの流れ

4. 速度問題の意味理解

速度問題の意味理解とは、算数問題解決システムにおいて、特殊問題に分類される速度問題を対象に、その問題文中の単語と数値より内容理解を行い、質問に対して適切な解を返すことができるということである。

次に速度問題とは、距離成分、時間成分、速度成分の三つの要素に関する情報より、質問された答えを求める問題である。その例として、「秒速4mで1分40秒間走ったときの道のりは?」のようなものが挙げられる。本研究で対象とする速度問題の種類は、「旅人算」、「通過算」、「流水算」、「他の速度問題」の4種類であり、前者3つを特殊な速度問題と位置づける。

4.1 処理の流れ

算数問題解決システムでは、まず入力文が算数問題であるかどうかを判断する。算数問題と判断されたら、特殊問題であるか一般問題であるかを判断するために特殊問題の判定を行う。そして、特殊問題、一般問題それぞれを正規化し、演算処理を行う。本稿では、特殊問題である速度問題に関する手法を提案する。図2に解決手法の流れを示す。

† 同志社大学大学院工学研究科

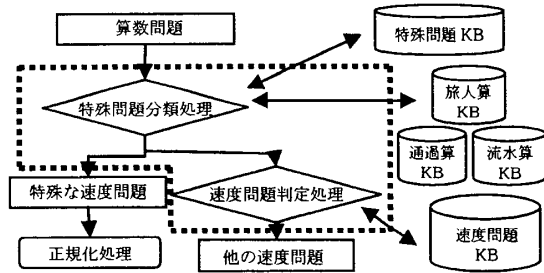


図2 速度問題解決の流れ

4.2 特殊問題分類処理

与えられた算数問題が特殊な速度問題であるかを判定する。先ほど述べた特殊な速度問題に関しては特殊問題分類処理において分類する。また、点線で囲まれている部分が図1における「特殊問題の判定」に当たる。

速度問題の中でも特殊な速度問題はその問題において特有の単語が使われるので、その単語を分類処理に用いる。そして、その単語を集めたものを知識として、旅人算知識ベース(5語)、通過算知識ベース(3語)、流水算知識ベース(7語)として作成した。

分類方法としては、質問文の自立語を抽出し、各知識ベースに格納されている単語との関連度を求める。その関連度を各種別の問題に存在する閾値と比較することにより、分類処理を行う。この処理により、質問文が「特殊な速度問題」か「それ以外」の二つに分類され、それぞれに対応した演算を行い、「それ以外」は速度問題判定処理にかけられる。

4.3 速度問題判定処理

速度問題判定処理とは、特殊問題分類処理において、特殊な速度問題と判定されなかった問題の中から、速度問題を判定し分類する処理である。ここでは、記事関連度計算方式^[4]を用いる。

本システムでは特殊な速度問題以外の速度問題はキーワードが存在していないので「他の速度問題」として扱った。そして、それを判定するための速度問題を問題文形式で10題格納した速度問題知識ベースを作成した。

まず、特殊問題分類処理によって特殊な速度問題と分類されなかった問題は、「他の速度問題」であるかどうかの判定を行う。具体的には、問題と知識ベースの間で記事関連度の計算を行い、その最高値が閾値以上かどうかを調べる。この閾値は、様々な問題をシステムにかけて得た実験結果から求めた値である。そして、分類された問題は演算を行い、解を求める作業に移る。

4.4 正規化処理

演算に必要な問題文中の様々な情報を「整理した形式的な表現に直すこと」を正規化と定義する。速度問題での正規化では距離成分、速度成分、時間成分の3つを対象に行う。正規化の手法としては各成分の単位に着目し、その単位の前の数値を抽出し、成分名(速度など)の3つで正規化を行う。正規化を行う際に、3つの成分を判断する知識として用いるのが正規化判断知識である。速度、距離、時間に関する単位をそれぞれ格納しておく。速度判断知識ベースに12語、距離判断知識ベースに4語、時間判断知識ベースに3語とそれぞれ格納されている。正規化を行った例を図3に示す。

時速 50 キロメートルのバイクが 10 分間に走る距離は何メートルか?
(速度, 50, キロメートル毎時)(時間, 10, 分間)(距離, 何, メートル)

図3 正規化の例

4.5 演算処理

正規化を行った後、分類した速度問題の種類別にメタ知識(数学における公式)に当てはめることにより、演算処理を行っていく。その一例を表1に示す。

表1 各速度問題の正規化項目とメタ知識の一例

	正規化項目数			メタ知識の一部
	距離	速度	時間	
通過算	2項目	1項目	1項目	距離①=速度×時間-距離②
他の速度問題	1項目	1項目	1項目	距離=速度×時間

4.6 評価

SPI問題集などから、速度問題を計60問(旅人算15問、通過算15問、流水算15問、他の速度問題15問)について分類処理、正規化処理、そして演算処理を行った。分類が正しく行われ、演算処理により解答が正解していれば、システムとして正解とする。その評価結果を表2に示す。

表2 評価結果

問題の種類	精度
旅人算	86.7%
通過算	73.3%
流水算	80.0%
他の速度問題	60.0%
全体	76.7%

表2からわかるように、全体で約76.7%ととても高い精度が得られた。処理できなかった問題の原因として、正規化項目の数の違いが挙げられる。表1に示したように、正規化項目の数をあらかじめ種類によって設定してあるので、それに当てはまらない問題は処理できない。こういった例外に対して、対応させるにはそれに応じた演算処理を構築するしかないと考えられる。

5. おわりに

本研究では、算数問題解決システムの特長の一つである速度問題の内容に着目し、速度問題における特徴を調べることで、多種多様な速度問題の意味理解を行う手法を提案した。これにより、速度問題に対する意味理解が可能となった。今後は、植木算などの他の特殊問題に関しても拡張および追加を行うことで、より幅広い算数問題解決の実現、そして、数量を含む自然な会話文を対象とした柔軟な判断の実現が可能になると思われる。

参考文献

[1] 眞鍋康人, 小島一秀, 渡部広一, 河岡司, “概念間の関連度やシソーラスを用いた概念ベースの自動精練法”, 同志社大学, 理工学研究報告, Vol.42, No.1, pp.9-20, 2001
 [2] 渡部広一, 河岡司, “常識的判断のための概念間の関連度評価モデル”, 自然言語処理, Vol.8, No.2, pp.39-54, 2001
 [3] 松山剛, 渡部広一, 河岡司, “常識知識を用いた算数問題解決システムの構築”, 信学技報, AI2006-55, pp.23-30, 2007
 [4] 倉田篤史, 渡部広一, 河岡司, “概念ベースと関連度計算を用いた記事関連度計算方式”, 情報処理学会研究報告, 2006-NL-171-(4), pp.19-24, 2006