104-OS-22a-1

知識体系の自律的構築と応用 Automatic Building of the Knowledge System and Applications

宮崎洋彰*1 Miyazaki Hiroaki

*1株式会社 オメガ・レゾン OMEGA RAISON CO., LTD.

The new system which builds the knowledge system autonomously, and performs various processing (conversations, solving problems) is presented.

1. はじめに

人工知能の知識獲得および知識を使った応用については国内外にて様々なアプローチにより研究が精力的に進められている. (参考文献[1]-[5]) その中で特に人工知能内に構築していく知識と現実との整合性をいかにして取るかについては重要な課題となっている. 著者は現実に対応する(=現実を表現する)情報(文)を逐次,入力することにより現実と整合した知識体系を自律的に構築し,構築した知識体系を使って様々な処理(会話,問題解決)を行うことができるシステムの開発を進めてきた.本稿ではシステムの概要および製作したプロトタイプの動作例について紹介する.

2. システムの概要

2.1 コンセプト

人間が知能を向上させていくプロセスとしては、最初は基本的な知識(事実、規則、正しい事等)を教えられ、次に入ってくる情報を記憶した知識と照らし合わせて評価し、考える。考えた結果を示し、示した結果が誤っていれば親および先生等から指摘を受け、正しい結果に修正していく。さらに進んでいくと問題に対して関連する知識を記憶した知識から検索し、処理を実施し、解決策を生成するという自律性を身に着けていく。このような「子供」が教育を受けて知能を向上させていくプロセスをアルゴリズムとして具体化する。

本システムでは人間の思考のベースである単語,数値および記号は既に特徴量であると捉えるところから出発し,特徴量の組合せ(概念,文,文章)から特徴量の組合せ(概念,文,文章)への遷移を人間が理解できる形で自律的に実施する.本システムの機能,性能向上の仕方は人間が「子供」を教育して能力および自律性を高めていくイメージに近い.最初は知識体系に基本的な情報(規則,事実,処理プログラム等)を入力し、情報間の関係を人間が設定する必要があるが,本システムは情報を入力していくにつれ情報と情報との接続関係の生成を学習していき,情報が互いに関係性を有した知識体系を自律的に構築していくようになる.後で入力する情報は知識体系を自律的に構築していくようになる.後で入力する情報は知識体系と照らし合わせて新規性,妥当性を評価しながら有用で安全な情報を知識として取り込んでいき知識体系を拡大していく.

人間は言語を用いて思考するので言語を構成する単語,数 値および記号をパターン化すれば,人間の思考はパターンの

連絡先:宮崎洋彰, 代表取締役社長,

横浜市戸塚区上倉田町 1372-2 A212, 045-862-1317, miyazakihiroaki2015@gmail.com

組合せで表現することができる.また思考過程はパターン間の 遷移で表現することができる.人間の思考は同じような状況下に おいては同じ思考過程を示す傾向があるため,関係の強い概 念間の接続関係を人間による設定および学習による自律的設 定により生成すると,ある状況に対する思考遷移をパターン間 の遷移によって表現することができる.またパターンの機能に能 動的な処理機能(情報検索,情報処理,処理結果の記録)を付 与することにより,人間の能動的な思考動作(条件付処理,逐 次処理,質問・命令への対応,有用な情報の記録等)もパターンおよびパターンの組合せで表現することができる.

2.2 機能およびプロトタイプ動作例

人間の思考および思考過程を模擬するシステムを実現する ため、パターンおよびパターンを格納する記録ユニットの機能を 以下のように設定した.

- ・入力した情報(文, 概念)の単語, 数値および記号を分析し, 記録器に記録する際, 主語, 主語の修飾, 述語, 述語の修 飾(5W1H) 数値, 記号および分析結果等の要素に整理して 格納し, 連想記憶的な情報検索を容易にする.
- ・プログラムの機能を包含する.(処理を担当するパターンが励起すると対応する処理プログラムが起動する.処理を担当するパターンをシーケンス的に実行することにより,複合的な処理を実行する.)処理の結果,新しく発生した情報は適切な検索キー(パターンの組合せ)により連想記憶的に記録し,他の処理からの情報検索を容易にする.
- ・パターンの励起により内蔵した処理を実行する. また接続関係のあるパターンを逐次, 励起し能動的な処理を実行する.
- ・パターンを格納している記録ユニットの接続情報を変更することによりパターンの励起シーケンスを変更する.

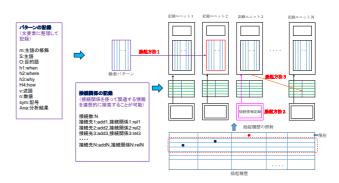


図1 パターンおよび記録ユニットの機能と励起方法

	1 E	回目の処理です。				
入力文 str_1 watasihasennseiidesu.						
入力文	入力文 str_2 わたしはせんせいです。					
NO	単語番号	単語 月	品詞番号	文要素番号	文要素	英単語
	0			1		I
1	1	は	11	1	主語	
2	2	先生	2	8	述語 (です)	a teacher
3	12	です	5	8	述語 (です)	am
4	5		21	12	文最後	
2 回目の処理です。						
入力文 str_1 watasihahouukagobukatudetenisuwooosieeteiimasu.						
入力文 str_2 わたしはほうかごぶかつでてにすをおしえています。						
NO	単語番号	単語 品	品詞番号	文要素番号	文要素	英単語
0	0	私	1	1	主語	I
1	1	は	11	1	主語	
2	13	放課後	25	3	H1(when)	after school
3	10	部活	2	4	H2(where)	the club
4	3	で	19	4	H2(where)	at
5	7	テニス	2	2	目的語	tennis
6	8	を	13	2	目的語	
7	9	教えています	- 7	7	述語 (他動詞)	teach
8	5		21	12	文最後	
3 回目の処理です。						
入力文 str_1 watasihatuyokusurutametenisuwokibisikuoosieeteiimasu.						
入力文 str_2 わたしはつよくするためてにすをきびしくおしえています。						
NO	単語番号	単語 品	品詞番号	文要素番号	文要素	英単語
0	0	私	1	1	主語	I
1	1	は	11	1	主語	
2	15	強くする	7	5	H3(why)	get strong
3	16	ため	24	5	H3(why)	in order to
4	7	テニス	2	2	目的語	tennis
5	8	を	13	2	目的語	
6	14	厳しく	26	6	H4(how)	hard
7	9	教えています	- 7	7	述語 (他動詞)	teach

図2 入力情報(語列)の構造を分析する実施例



図3 入力情報の記録(知識体系の構築)例

図 1 にパターンおよび記録ユニットの機能と励起方法について示している. 主な励起方法は下記である.

- (1)検索用のパターンを各パターンと照合し、相関の大きいパターンを励起する.
- (2)現時点から設定した期間の過去の励起履歴を各パターンの接続情報記録部に記録した他のパターンとの接続情報と照合し、相関の大きいパターンを励起する.
- (3)接続関係のあるパターンを励起する. 接続関係が複数ある場合は設定した関係のパターンだけを励起することができる.



図4 入力情報と記録情報の比較例

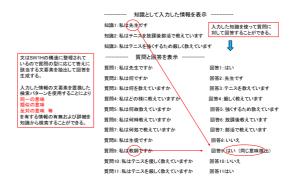


図 5 知識体系を利用した質問への回答例 1

パターンの最少単位は語,数および記号に対応するものである. 語が時系列的に発生することにより単語に対応するパターンが励起する. 数および少数点に対応するパターンが時系列的に発生することにより数値を検出する. 記号および記号の組合せは適宜,対応する演算子等に識別し,識別結果を記録する領域に記録する. 語が時系列的に発生することにより単語が識別され,単語に対応する品詞の情報を使って文要素を識別していく. 主部は主語と主語を修飾する部分に識別し,述部は述語と述語を修飾する部分に識別する. 述語を修飾する部分は(いつ,どごで,何を,どのように,何故)が識別される.

単語,数,数式,記号からパターンへの変換は識別できるも のであれば特に形式に制約は無い. またパターンの組合せも, パターンである. 本システムのプロトタイプでは各単語に識別番 号を割り当てることにより実施した. パターンを番号で表現する ことによりパターン間の処理を演算で実施できる長所がある. (例えば「私」という単語の単語識別番号は「0」, 助詞「は」の単 語識別番号は「1」、「先生」「2」、「です」「12」、・・・「人工知能」 「169」,「職業」「170」・・・等々) 語の励起履歴が各単語の接続 情報記録部に参照され照合が行われる. 単語に対応する語列 および一致した語数を検出すると該当の単語が励起する. 単語 が検出されると検出した単語の語列を除いた励起履歴を使って 次の単語の検出を行う.この時,次の単語を検出できれば,先 の単語の識別(区切り方)は適切であったと判断できる.一方, 次の単語が検出できず不整合が生じた場合は先の単語の識別 (区切り方)は適切でなかったと判断し別の単語候補に対し整 合性を確認する. (単語の区切りによる整合性を逐次実施)

図 2 は入力情報の構造の分析例を示したものである. 語列 から単語, 単語識別番号, 品詞, 品詞識別番号, 文要素および 文の構造が分析され分析結果を使うことにより文要素を検索し 易い構造に整理した記録が可能となる. 図 3 は構造分析した情 報を検索し易い構造に整理して記録している様子を示している.

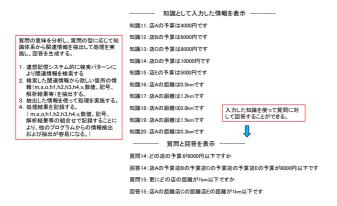


図6 知識体系を利用した質問への回答例2



図7 知識体系を利用した質問への回答(詳細処理)

 $m1\sim m4$ は主語の修飾部、s1,s2 は主語、o1,o2 は目的語、h11,h2 は when、h21,h22 は where、h31,h32 は why、h41,h42 は how、v1,v2 は述語に対応する文要素の単語識別番号および数値を格納している。入力情報の単語、記号、数字、式はパターン(識別番号)に変換し、主語、主語の修飾、述語、述語の修飾(何を、いつ、どこで、なぜ、どのように)等の要素を検索し易い構造に整理して記録している。

このように入力した文を構造化して記録すると文と文の比較を 文の構成要素毎に実施できるため、非常に有益である。例えば 完全に同じ文を検索する場合は全ての文要素が一致するパタ 一ンを検索する。一部が一致すれば良い場合は一致させる箇 所を組み合わせたパターンを検索パターンとして使い、該当す るパターンを検索すれば良い。このように文要素を構造化する ことにより情報を連想記憶的に記録および検索ができるようにな る。

図 4 は入力した情報のパターンと記録領域にある情報を文要素毎に比較し、文要素のどの箇所が一致しているかを示したものである. 入力した情報と記録している情報とのパターンを文要素毎に比較することにより、完全一致する情報または部分的に一致がする情報が記録されているか否か識別でき、入力情報の新規性および関連情報との関連性が評価できる. また入力情報の文要素に関し、別の関係性のものに置換して検索することにより、様々な情報を記録した情報から抽出することができる. 例えば入力情報のある文要素を類似の単語のパターンに置き換えることにより、類似の意味の文を検索することができる. また

入力情報のある文要素と反対の意味の単語のパターンに置き 換えて検索することにより入力した文と反対の意味を有した情 報の有無を確認することができる. このことを応用すると入力し た情報が規則およびモラルに則したものか否かを判断すること が可能となる. 知識体系に知識を記録する際に, 規則およびモ ラルに関する情報は、その識別結果(規則およびモラル)も付帯 情報として合わせて記録しておく. 例えば入力した情報と記録 した情報を照合し不整合(反対の単語検出)が検出され、記録 した情報の付帯情報が(規則およびモラル)である場合,入力し た情報は(規則およびモラル)に反していると識別することがで きる. 図 3 において文番号 1~3 に図 2 の文(知識)1~3 に対 応する情報が記録されていることが分かる. 図 4 は入力情報と 記録情報を比較している動作について示したものである. (a)の 比較では第4番目に入力した文として第2番目に入力した文と 同一の文(私はテニスを放課後部活で教えています)を入力し ている. 第4番目に入力した文は文要素毎に第1~3の文と比 較され文要素毎に一致するか否かが確認される. 入力した文と 完全に一致した場合, total の欄に '1' が表示される. 本例 では文番号2のtotal欄に'1'が表示されており入力した文 は既に記録している文番号 2 の文と一致していることが識別さ れている. (b)の比較では知識と反対の意味を有する文(私は 生徒です)を入力した場合の比較結果である.この入力文は文 番号 1 の文(私は先生です)と整合しない情報なので文要素お よび total の欄に反対を示す'-1'が表示される.

本システムは質問から検索パターンを生成し質問に対する回 答を生成することも可能である. 図 5 では質問 1~11 に対し回 答が正しく生成されていることが分かる. 単語と単語の関係性 (同じ意味, 反対の意味)を各単語に対応する記録ユニット間で 定義しているので、質問として同じ意味あるいは反対の意味を 検出すると回答に反映するようにしている. (ある情報を評価す る時, その情報だけでなく類似の単語, 反対の単語を使った関 連情報による照合も実施することで実現)質問9では「私は教 師ですか」という内容になっている. 事前に設定した知識として は「先生」であり、「教師」という知識を明示的に与えていないが 同じ意味を表す「教師」に対しては「はい」と正しく回答している. このように単なる形式的な照合ではなく記録している情報と意味 的に比較し評価することができる. 図 6 では数字を含む文を分 析し処理を実施している例である. 図7は質問 14 を構文分析し 文要素に識別するとともに質問の意味を分析し回答を生成する 処理について示している. 文要素のパターン(どの**の○○が **)から質問の種類を検出している. 質問 14 では〇〇は「予 算」が対応する. 知識体系に関連する情報が記録されているか 検索する.この時に使用する検索パターンとしては例えば「**」 「の」「予算」「は」「?」「です」というパターンを使用すれば良い.

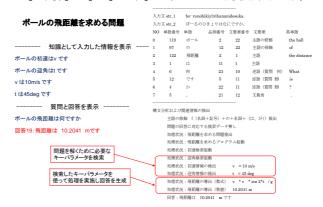


図8 問題の型を検出し自律的に解を生成する例1

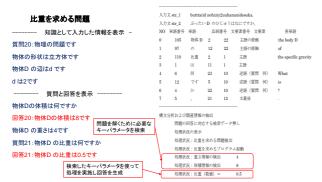


図9 問題の型を検出し自律的に解を生成する例2

記録ユニットは情報を構造化して記録するとともに情報の処理の仕方も記録する. またプログラムの機能を包含する. これらの機能の実施例について次に説明する.

図 8 は物理の問題(質点の飛距離)である. ボールを初速: V m/s, 迎角: t deg で投げた場合の飛距離について求める問題である. システムへの知識としては①ボールの初速は V です. ② ボールの迎角は t です. ③V は 10 m/s です. ④t は 45deg です. が入力されている. 次にシステムへの質問として Q:ボールの飛距離は何ですか?という質問を入力している. この時,システムの処理状況を図 8 に示す. システムはボールの飛距離に関するデータが既に存在する否かの確認を第 1 ステップとして実施している. 飛距離に関するデータが存在していない事を確認すると,飛距離を導出するプログラムを起動し,ボールの初速および迎角に関するデータの検索を行っている. ボールの初速はV(記号)で数値は 10 m/s であること,ボールの迎角は t(記号)で数値は 45 deg であることを知識から検索している. 次のステップで飛距離の数式は $V*V*\sin2t$ /g であり,数値は 10.2041 m であることを回答として正しく出力している.

図 9 は物理(比重)の問題である.システムは最初に質問に対する回答(物体 D の比重)があるか知識体系を検索している.検索の結果,該当のデータが無いため,比重を求めるプログラムを起動している.本システムは物体 D の体積および重さのデータを検索している.物体の形状に関する情報(立方体)および寸法情報から体積を求め,物体の質量情報を検索することにより比重を求めている.

3. 本システムの応用

本システムは入力した情報を文要素(主語,主語の修飾,述 語, 述語の修飾(5W1H), 数値, 記号) に整理し構造化して記録 する. 情報の構成要素である単語, 数値および記号を識別可 能なパターンに変換することにより、情報はパターンを組み合わ せた構造に変換され知識体系を構成することが可能となる. 知 識体系はパターンの組合せにより情報を抽出することができる 記録システム(所謂, 連想記憶システム)となっている. 情報の 記録および検索を適切にパターンの組合せることで実施するこ とができるので、ある処理プログラムで生成された情報を他の処 理プログラムで使用するということが柔軟に実施できる. これによ り処理プログラムのユニット化が可能となる. 図 10 は情報を連 想記録的に記録および検索を実施することにより柔軟に処理が 実施できる動作について示している. 通常のプログラミング手法 では、あるプログラムで生成した情報を他のプログラムで使用す る場合は受け渡しの変数を厳密に定義する必要があるが、 逐次 新しい情報が生成するような状況には対応できない. 人間が思 考する場合, 思考の過程で生成した情報を思考のワーキング領 域に一時的に記録し、次の思考過程で利用すると想定される. このような思考間の遷移を本システムの連想記憶システム的な情報の記録および検索を活用することにより、各処理プログラム間の遷移により実現することが可能となる. 処理プログラム(処理を実行する記録ユニット)の構成を以下に示す.

- ○問題識別部:問題の種類を識別し、対応する処理を起動する. ○主要パラメータ検索部:問題を解決するために必要なキーパラメータを検索する.キーパラメータが検索できない場合はキーパラメータを生成するプログラムを起動する.
- ○処理部:主要パラメータ検索部において獲得したキーパラメータを使用して処理を行う. (処理は数式処理,言語を使用した処理も含む. 各問題の型に応じて適宜,設定する箇所である.) ○記録部:処理部において生成した処理結果を新しい知識として知識体系に記録する. 記録は連想記憶的に行う.

各種問題の処理に対応できるよう処理プログラムのライブラリを充実させることにより、数々の複合的で複雑な問題を処理するシステムを実現することが可能である.

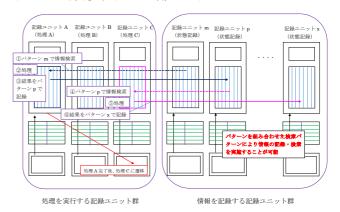


図 10 連想記憶的な情報の検索・記録および処理の実施

4. おわりに

本論文に掲載した内容は参考文献に示した特許にてカバーされています。著者は本人工知能に関する技術、手法が広く人工知能、知能ロボット分野に使われることを希望しており、特許の使用に関して実施許諾の用意があります。使用の意向または興味を持たれた方はご連絡お願いします。

参考文献

[鈴木 16]鈴木: 実体ベースの概念からプロセスベースの概念へ,

人工知能学会誌, Vol.31, No1, pp52-58

[谷口 16]谷口:記号創発問題

人工知能学会誌, Vol.31, No1, pp74-81

[Clark 13]Peter Clark, Phil Harrison. A Study of the Knowledge Base Requirements for Passing an Elementary Science Test. AKBC'13, October 27-28 2013, San Francisco

[Jansen 13]Peter Jansen, Mihai Surdeanu, Peter Clark.
Discourse Complements Lexical Semantics for Non-factoid
Answer Reranking. Allen Institute for Artificial Intelligence.

[Luis 06] Jose-Luis Ambite, Vinary K. Chaudhri, Richard Fikes, Sunil Mishra, Maria Muslea, Thomas Uribe, Guizhen Yang. Design and Implementation of the CALO Query Manager. AAAI 2006.

 [宮崎 14]宮崎 自律型思考パターン生成機
 特許第 5737641 号

 [宮崎 14]宮崎 自律型知識体系構築機
 特許第 5821141 号

[宮崎 15]宮崎 暴走しない人工知能装置 特許第 5854392 号