

1. はじめに

言語処理系は、認知系や思考系、学習系と深く関わります。言葉を手繰るとは、そもそも基盤となる情報処理能力が最高に発揮される所にあります。言語処理が実現できるならば、その技術を他の情報処理に横展開し、直ちに構築できるでしょう。

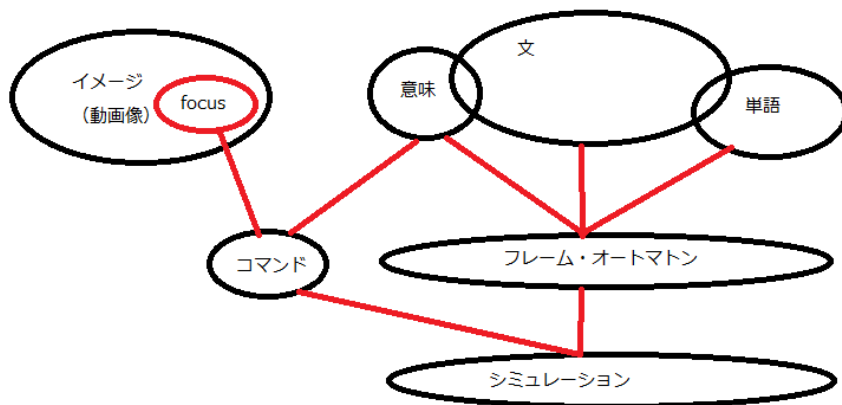
言語処理系は、言葉の獲得、言葉の認知、意味処理・思考、言葉の表出とからなります。基盤となるデータエリアは、コンセプト（コマンド、単語）とそれらのセット、グラフ、イメージであるマップ、コンセプトの関係を示すニューロンとフレーム・オートマトンがあります。そうして、それらデータエリアの上にオントロジーデータ、オントロジープロセスが配置されて言語処理系は構築されます。言語処理系は自己組織化していくものとして構築され、どんな言語でも自律して獲得し、操作していけるように造られます。

2. 言語処理系の基本的考え方

言語の場は生活空間であり、これは動画像や音声などのイメージデータ空間でもあります。そうしてそれは、身体、感情、思考などの処理の対象であるコンセプトに関する意味世界と人工知能により繋がるものです。

言葉は単語と文法により文を構成し、文がシナリオやレトリックによって束にされて文章になるというスキーマに則って紡がれるものです。

言葉を獲得し、利用するのは、言葉の持つ構造を推論して行く過程が必要になります。その過程のもとになる構造データはフレームとオートマトンです。



知識は断片が繋ぎ合ったものとして表現するのが単純で強力なことから、フレームというタイル張りの情報処理がなされます。また、言語は 1 次元で時間を追って情報を並べるということで、オートマトンで表現するとスッキリします。

また、人工知能が生活する中で、生活の場から得られるデータは一次元、2次元、3次元のマップとして表現できるイメージというものです。イメージは人工知能の持つ情報の全てに於いて基盤となるものです。

イメージは階層構造をしています。そうして、**Focus** によって、イメージから必要な部分を切り取って、コンセプトと対応付けします。街を歩いているイメージは、街の道路や建物があって、イメージの中の点に道路のイメージ、建物のイメージが添付されていて、人のイメージが道路のイメージの中の点に人動画イメージとして添付されていて、道路上の人の点が道路イメージ上を移動していくと表現されます。木々のイメージも街に添付され、道路のイメージにも木々のイメージが添付されることもあるでしょう。**Focus** は、街のイメージのどこに視点があって、視界が設定されているかを管理し、解析処理（認知）の枠組を与えます。

動画はチャンク構造の集まりで、シミュレーションの為に利用されます。人物の街を見回す回転ということも、各画像イメージを順を追って提示することで、一定の速さで変化していきます。これは、チャンク毎に評価処理をするためです。シミュレーションの結果を刻々と評価していけるようにします。実際の速度や加速度は動画の属性として記号（コンセプト）で表現され、動画画像イメージに添付されています。

フォーカスには、ズーム イン、ズーム アウトとパンとチルトの機能があります。より上位の視点から見るのがズーム アウトで、細部にフォーカスするのがズーム イン、近傍に視点を移すのがパンとかチルトです。そのフォーカスの移動で、まだコマンドがイメージに添付されていない時には、イメージを新しく解析してコマンド表現を得ることが（思考処理）が必要になります。この解析を思考過程とする理由は、目標を設定し、解析プロセスを選択し、結果を評価していくからです。

3. 認知オントロジー

基本的認知オントロジーは、

- ・ 存在、空（真、偽）
- ・ メリット、デメリット
- ・ 肯定、否定
- ・ 包含関係
- ・ 順序関係（比較関係）
- ・ 原因、結果
- ・ 始点、中間点、終点
- ・ 特徴点
- ・ 配置（方向）
- ・ 一対一対応
- ・ 強さ（大きさ）
- ・ 近さ（遠さ）

であると思えます。そこで、日本語の助詞、助動詞を調べてみました。助詞の働きとしては、

- ・ 主語（主役）
- ・ 修飾関係
- ・ 属性の名詞化
- ・ 作用の結果、場所、時間、到着する所、行動の目的、受け身とか使役の相手
- ・ 方向、帰着する場所
- ・ 共同、作用の様子、引用
- ・ 動作の起点
- ・ 比較、限界
- ・ 手段、材料、場所、原因、理由
- ・ 仮定、追加的列挙
- ・ 一定条件のもとで何時もある事柄が起きる様子
- ・ 場合
- ・ 結果
- ・ 成り行き
- ・ 前と相反する関係
- ・ 確定、対比
- ・ 逆接
- ・ 連接
- ・ 中止

助動詞では、

- ・使役
- ・受け身
- ・可能、自発、尊敬
- ・希望
- ・打ち消し
- ・推量
- ・様態
- ・不確実
- ・断定
- ・例示
- ・丁寧
- ・過去
- ・完了
- ・意志

こうしてくると、言語の枠組を規定する、オントロジーが見えてくる気がします。これらは、他の言語でも顕著な項目なのか、知りたい気もします。

人工知能設計者が設定したオントロジーと、人工知能が実行時に獲得していくオントロジーが矛盾無く整合性を保てるようにする特別な機構が必要です。オントロジー記号は他のオントロジーやコンセプトに連想関係を持って、その働きが定義されます。だから、その連想関係がグラフとして、同型であれば同じオントロジーとして、人工知能設計者のオントロジー記号を採用するように常に監視することが肝要ということになります。

なお、画像解析のオントロジーとして次の物が上げられます。

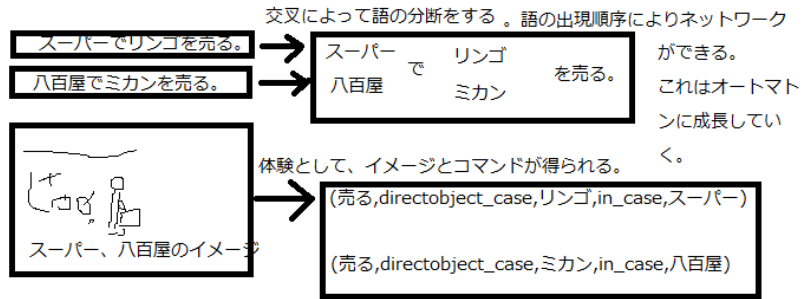
- ・幅、始点、終点、中間点
- ・直線性
- ・曲率
- ・角、角度、スピン
- ・変曲点
- ・繰り返し、包含関係
- ・オブジェクト
- ・属性
- ・格
- ・主人公、話し相手、操作の対象
- ・時間、空間、事象

文の要素の意味を作るのはイメージの解析による物です。そのイメージの解析の目的となる特徴というものをあらかじめ設定しておくことは、推奨されますが、必須ではありません。イメージ画像の重なりを沢山重ねると、重なり濃淡が出現します。この濃淡の様子を特徴量として設定していけば良いのです。その特徴量を解析するように、解析コマンドを生成していくと、自己組織化的にイメージ解析プロセス群をそろえる事となります。

解析にはポテンシャル法が有効になることがあります。画像解析には各点からの距離によって小さくなる対称な重み関数を画像全体に貼ることをして、その重なりを解析していきます。言語空間ですと、コンセプトセット内に何かのメトリックで距離を導入して位相空間にし、その位相にポテンシャル関数を貼り付けて、解析していきます。

4. 交叉法とフレーム・オートマトン・単語選択

言語の基本的な構造は交叉法で学習して行き、その結果はフレームとオートマトンによって表現されるデータになっていきます。



コマンドとイメージの解析から、意味素性とか格とか、基本的な意味構造が得られています。これと表層表現である文は、交叉によって、語を得られると同時にコマンドとの交叉によって、語への意味の付加ができるようになっていきます。

コマンドの構成についてですが、行動の枠組は動作であり、動詞で表現され、行動の枠組中の部品である登場人物や行動の対象などは名詞になり、動詞や名詞の属性が副詞や形容詞になり、行動の枠組と部品の関係が格になるというようにイメージを解析して得ていきます。

交叉法では、区切り記号の発見が有効です。区切り記号でデータ群を整列させていくのです。日本語だと区切り記号は助詞、助動詞で、英語では冠詞や前置詞です。音声では、拍とかアクセントなどのリズムも大切な整列処理の手がかりになります。

なお、イメージとコンセプトの間の対応にも交叉法は有効です。「スーパーにみかんとリンゴが売られている」イメージで、コマンドに「(売る,directobject_case,リンゴ,in_case,スーパー)」と「リンゴ」のみを指定していたとして、このとき、みかんとリンゴのイメージの両方が「リンゴ」に対応しながら、全体としてイメージとコマンド対が作られます。次に、「スーパーにマンゴーとリンゴが売られている」イメージを提示して、コマンド「(売る,directobject_case,リンゴ,in_case,スーパー)」が与えられたとすると、両方のイメージで共通なのが「リンゴ」ですから、「みかん」のイメージは「リンゴ」と異なること、それと、リンゴ以外のコマンドの中身を学習済みならば、「リンゴ」のイメージを特定できます。

5. ディープラーニングと言語

大局的な言語の持つ構造はディープラーニングで学習していくことになります。結果のデータはパーセプトロンのネットワークになるということです。

ディープラーニングの基盤となる処理は、共起に基づく連想をコンセプト間に張ることです。生活しながら、人工知能が自律的に連想関係を作っていくのです。それは、「長野」ですと、「地理的」関連で行くと、「善光寺」とか、「県庁」とか、「長野駅」があり、「文化」ですと「おやき」とか、「獅子舞」とか、「節分の種まき」とかになっていきます。その中でエピソードが串刺しになってまた一つの連想のベースとなる事象になっていきます。後述する「意図」の推定に使われる連想は、「主体」の「行動」が「何を期待してのものか」という、人間的な因果関係を対象にしたものです。このように、連想というものは枠と部品群というセットが幾重にも重なったものとして表現できると言えます。

文章を読んで、どのシーンのことを語っているかを問うことを考えてみたいと思います。

(1) 列車の中のシーンの特徴

- ・乗り降りの手続きがある
- ・椅子がある
- ・人がいる
- ・窓があって外の景色を楽しめる
- ・外の景色は流れていく

など

(2) デパートの中のシーンの特徴

- ・建物のなかの景色がある
- ・人が居る
- ・椅子は殆ど無い
- ・売り場がいっぱいあり、商品が沢山並べられている
- ・外は見えない場合が多い

など

(3) レストランの中のシーンの特徴

- ・建物の中の景色がある
- ・椅子がある
- ・テーブルがある
- ・食事をする
- ・人がいる

など

こんな中で、「窓の外の景色の流れる様をみていた」ですと、かなりの確信度で、「列車の中」を答えにします。「椅子には人が座っていた」ですと、「列車の中」か「レストランの中」か迷います。そんな推論をしていくのにディープラーニングでの学習は有効です。セット処理の中の、部品群による枠の推定ですね、コウユウ事に長けたのがディープラーニングであるということです。

話者の意図をくむ方法についても、ディープラーニングは有効です。コンセプトセットが何を意図していくかのコンセプトを連想させるのですね。場合によっては、**And/or/not Set** 処理がディープラーニングの入力ノードセットの選択の為に必要になります。

また、話者の意図は、コンセプトセットから生成されるイメージでシミュレーションしていき、途中経過とかシミュレーション結果とかのコンセプトセットに対してディープラーニングの入力ノードセットとしての割り付けによって、意図をくむことをしなくては成らない場合もあります。

例えば、

「チョコレートあげる」は、「付き合っていて悪い気はしない」という意味だとして、この2つの事象がコンセプトレベルで連想関係を持っているとしてよいでしょう。しかし、「チョコレートが余ったから関心をかきたい」、「ホワイトデーのお返しが楽しみ」という意味かも知れません。「チョコレートをあげる」はいくつかの事象と「意図」という情報クラス（セットとして管理されるコンセプト群）内で連想を持つのですね。また、チョコレートの材料を貰ったお返しで、「チョコレートをあげる」かも知れません。エピソードのあるステージでの事象群を想定して分かる、「意図」かも知れないのです。ディープラーニングでなく、事象コンセプト（エピソード）の場における連想であり、それが「意図」というクラスの連想のものを探すということが意図の理解の処理であるとも言えます。幾つかの機構が重なったものが意図理解機構ということですね。

ディープラーニングとセット空間。この両方が同時に、情報処理に重要であると言えます。

情報が、枠とその部品群という基本構造を持っていることから、情報処理の基盤的な事柄は、枠をセットのオーナーとし、部品群をセットのメンバーとする操作によって実現できることになります。交叉法もディープラーニング（重み付き投票）も、ベイジアンネットワークも大方の仕組みは、セットの操作を基本とすることです。これに付け加えて、連想も発想も、このセットが重要になることも付記します。

パーセプトロンの入力層とか出力層のノードセットを決定するのも、このセットの重なりを **and** で取るか、**or** で取るか、**not** で取るかという集合論的な操作です。

一つのコンセプトには相互に無関係な無数の種類のコンセプトが連想しています。「長野」だと、「受験の思い出」、「山」、「戸隠」、「善光寺」、「千曲川」、「除夜の鐘」、「ソバ」・・・とうとう。そのうち、「年末だね」というような話題が登ったとします。その時、連想する・・・発想する・・・のは、「除夜の鐘」とか、「除夜の鐘」から「善光寺」へと連想が連鎖します。「ソバ」も「年越しソバ」の連想から発火するかもしれません。このように、目的の物だけを効率よく連想するには、「年末」との連想のセットと「長野」の連想のセットの重なった部分だけを検索するようにすると良いことが分かります。このように、発想とか連想とかにもセットが強力だと分かります。と同時に、実際にコンセプトが発火するのではなく、潜在的に発火しているという状況が重要になります。潜在発火の重みは、セットの重なるの重みであることです。

6. シナリオ、レトリック、文法、単語選択

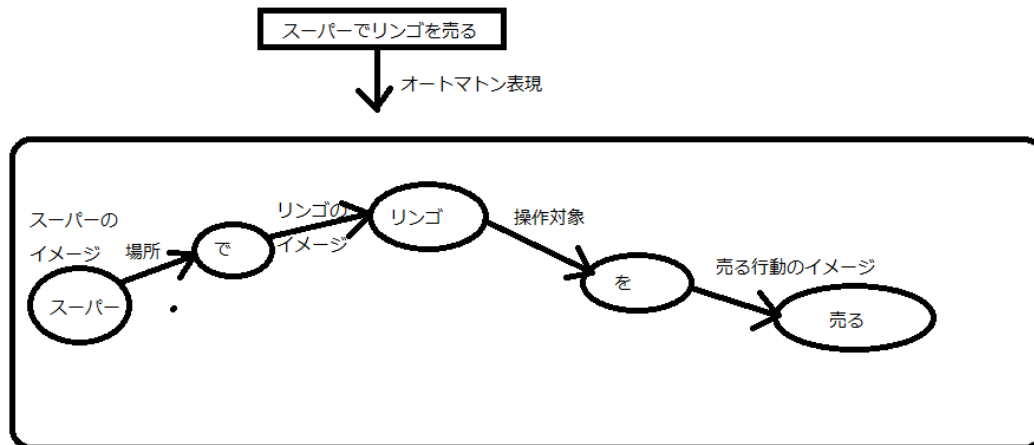
シナリオはチャンクというまとまりの入れ子と連鎖によって構成されます。ここでいうチャンクは、2つの文の対で、次のものがあります。

- (1) 時間シーケンス
- (2) 空間シーケンス
- (3) 包含関係
- (4) 原因と結果
- (5) 起点と中間点群と終点
- (6) 飛躍と着地
- (7) 提示と例題・説明
- (8) 強調、肯定、否定、言い換え
- (9) 問いと答え
- (10) 挨拶と本文

これらのカテゴリーで評価しながら文が配置されていきます。

「私は昨日長野駅へ行った。晴れていたし、本も探してみたかったです」という文では、主題の提示と、その説明という文の対で文章を作っています。文章はこのようなリズムをもって語られていくのが常です。

レトリックと文法はフレームとオートマトンによって管理され実現されています。



このオートマトンは、スーパーのイメージからそのなかの「売る」行為のイメージに **Focus** していき、そのフォーカスした行為の意味をコマンドとして認知していることを基盤として造られます。認知はイメージの動画像認識によって得ている物とします。

単語選択は、コマンド内の同じ名詞、属性詞から異なる活用の単語を対応しているときに、造られます。これもオートマトンの遷移によって表現されます。

格と名詞・属性詞の塊は、コマンドの性格から根本的に、付加や削除が自在です。このため、オートマトンが幾つかの塊に分断できると考えた方がスッキリします。文の中に文がある構文とかに対応するためにもこの考えは必須になります。このオートマトンの分断した個々の要素をフレームとして特別な考慮を払うこととします。

言語認識・表出でキーとなる技術である、オートマトンプールをどう利用していくかということを考えてみます。状態をコンセプトとか、単語・機能語とします。遷移記号は格とか意味ですね。状態も遷移記号も含意関係の解決は行われているとします。このようにオートマトンを造っていくと、文解析では、ステートを埋め込んでいく（プール内にある無数にあるオートマトンで、平行処理で、その単語が埋め込まれていきます）そうして、最初から最

後までのステートが埋め込まれたものが正しい解釈と判断して、その時の意味と格を遷移記号から得ます。それで、深層表現であるコマンドが構成できます。逆に文生成では、コマンドで、格と意味が与えられます。そこで、遷移記号を決定して、プール内の無数のオートマトンで、その遷移があるところに埋め込んでいきます。で、最初から最後まで遷移が埋め込まれた物が正しい文であると判断して、その時の単語を決定し、文としていきます。

思考に言語が中心的な役割を担うのはなぜか。シナリオ機能が思考にも重要で、単に直感的な思考では言語は必要ない物です。このシナリオによって、思考の展開が効率的にできるようになるのです。数学や物理学ですと数式を中心にして推論を展開します。数式やロジックというのはシナリオであり、言語システムが必要なものなのです。

7. 文の構造

哲学書とか、長い文章とかを読むときに、文単位では読めても、全体として一体何を言わんとしているのかと迷子になってしまうこと多。ぼつんぼつんと島のように分かるところがあるというようなこと。認知心理学でも、そんな例文がいくつかあります。文を一つ一つは分かるのですが、全体のイメージが構築できない。後で、こうゆう状況のものでよとか、枠組のシーンを教えられると、一気に理解できる。・・・そっか・・・と笑いが漏れる。これは、認知にも枠組・・・シナリオというものが重要であることを示しています。「椅子に座った」といえば、映画館かな、電車の中かなとか、会議かなという推論が働きます。しかし、それに続いて、「水が沢山あって」なんて文が続くと途端に意味不明の迷子になっていく。「ボートレースの状況さ」ということであれば、合点がいくことに。

つまり、シナリオというか、枠組というものの認知における重要さが見て取れるわけです。長文読解もシナリオを自律学習できるシステムにすれば、難なくこなせるということです。

枠と部品との関係といえば、「国境の長いトンネルを抜けると雪国だった」という良い例で説明ができます。この文だけでは、人が県境近くのスキー場にスキーにいったというような枠組で認識するでしょう。それがスキーが好きな人の世界の枠組に近いからです。しかし、旅行を良くする人からは、「国境」、「トンネル」が提示された状況から、「列車」しかも、文章が書かれた時代から「蒸気機関車」で、「雪国」・・・といえば「新潟」を推論します。以上のことは、部品である「国境」、「トンネル」、「雪国」がどの枠に納まるかということ推定する問題であるということです。

国語の入試問題を読みながら、文章を理解するとは本当にどんなことなのかということを考えていました。単に、日本語を中間表現で置き換えるだけで済まし、問題提示に応じて、その中間表現を解析していくという方式で良いのかという問いですね。問題に答えるだけならばそれでも良いと思えることもありますが、やはり、人間的に分かったと言える物にしたい。やはり、枠と部品群という関係ですね。このとらえ方を入れ子にして、知識として文章内容の解釈を構築して行きたい。これが、文章を本当に分かったと言える状態なのではないか。そのとき、枠は時間と場所とテーマコンセプトですね。そうして、文章には中心的なアクターとして、登場人物・登場オブジェクトと主役があります。これらは部品なわけで、その部品が互いにやりとりを行う。部品には属性がある。部品は、時間と場所によって変化していく。そんな情報がオブジェクトには添付されていく。このように、明確に枠と部品群と部品群の持つ情報という形で、文章の内容を把握しておけば、様々な連想処理にも簡単に対応できるし、(when,where,who,what,why,how)として把握し直すことも、リーズニングをすることも簡単です。

8. 言語処理系の構造

言語処理系の領野としてどういうものを構築すれば良いかを今までの議論から考えてみます。

(1) データエリア

- ・シナリオプール
- ・フレーム・オートマトンプール
- ・単語プール
- ・多様体的知識データエリア

(2) プロセスエリア

- ・イメージからコマンドを解析によって得るプロセス群
- ・コマンドからイメージを合成していくプロセス群
- ・データ検索プロセス群
- ・連想処理・ディープラーニング・交叉法プロセス

(3) 知識ベース

- ・常識
- ・エピソード
- ・専門知識

(4) 関連領野

- ・認知システム
- ・シミュレーションシステム
- ・思考システム
- ・行動システム
- ・ミラーニューロンシステム
- ・評価処理システム

基本データ形式は、イメージであるマップと、連想関係と遷移関係であるグラフ、そして枠と部品の関係であるセットと知識要素であるコマンド、(key,value)があります。全ての知識はこれらの組み合わせで構成されます。その知識の要素はコンセプトとして、マップ、グラフ、セット、コマンド、(key,value)を包む物になっています。

そうして、グラフとコマンド・(key,value)の組み合わせから、オートマトン、単語、パーセプトロンが生成されます。それらが更に集まって、プールや知識ベースに成っていきます。

知識ベースは、常識とエピソードと専門知識などが埋め込まれていき、利用されます。そうして、思考システム、行動システム、評価システムなどのシステムが実現されます。

9. おわりに

意味の深層表現であるコマンドをイメージから得、一方で表層表現である文を音声やテキストを得て、単語や文法、レトリック、シナリオという言語能力を獲得していく機構を説明してきました。それには、文と意味をオートマトンという表層表現と深層表現の混合した表現を用いて行うことを示しました。フレーム・オートマトンという考え方によって、言語の獲得から利用まで、見通しよく実現できることが示せたと考えます。そのフレーム・オートマトンプールを造るために交叉法を使いました。交叉法は言語要素の枠と部品群でつくるセットの操作でもあります。

交叉法とオートマトンが言語処理の基盤技術であることが分かっていただけたと思います。それで、交叉法とディープラーニングの基盤である **and/or/not Set** 処理というものによって、どんな言語でも獲得可能であることが重要です。言語を利用して行くには、コマンドなどの意味要素を保存するコンセプトの間の連想関係も重要になります。ディープラーニングについても触れました。

ディープラーニングは強力な推論能力を発揮します。しかし、それでも精密な推論や未知の状況への論理だった推論が必要な場合には、言語を司るシナリオ、文法、単語選択、イメージ解析プロセスなど、オートマトンやノイマン型コンピュータが必要になります。脳でもそんな手続き的な神経回路網があることでもプロセスの必要性が分かります。

以上説明してきた技術を体現した人工知能が、生活しながら、世界の言語を学習して後世に残していく・・・大きな目標の一つでもあります。

映像から言葉へ、言葉から映像へ。そうして思考があつて、万能の言語処理が実現できることが示せたとと思います。