

本来のテーマである自然言語処理についてシナリオをもって分析してみました。今回はそれを記していきたいと思えます。

その前に、「発火」機能ですが、促通と抑制の2つがあるのです。これは神経の連絡が促通と抑制の2つ有ることがヒントだったのですが、シナリオ分析していてどうしても、この2種の存在が必要になりました。促通と抑制の2つは、2つのオブジェクトの間の関連を表すオブジェクトにプラス、マイナスの2つを持つことで実現します。関連の強さにプラスとマイナスがあるということですね。

1. 日本語の解析の目標

日本語処理では、ステージ、シーンを認識し、アクターを把握し、アクター間の関係とアクターがどうゆう状況に有るかを把握することが、文章解析の目標です。状況として把握するのは次の物です。

(1) ステージ

会話がもたれている、時間と場所を把握することは日本語処理の基本です。ステージはその基本的な情報を表します。

・アクターの has_a, is_a 関係

登場人物(アクター)の間の関係を解析することが日本語処理の最大の目標でしょう。その内で特に has_a 関係と is_a 関係がステージと大きく関係します。例えば「私は汽車に乗った」ですと、汽車が私を has することになり、あとの話しの設定が汽車の中の事になる可能性が高いです。汽車の中がステージになるということです。

・現行ステージの把握

has_a 関係、is_a 関係の内、大きな枠組みのアクターが現行のステージになるのです。

(2) シーン

シーンは基本的に、登場人物が一定の場面ということになります。シーン解析はアクターの把握が基本になります。

・全てのアクターの把握

物も概念も名詞で現れるとアクターとして扱います。

・人物アクターの把握

人間は特に重要なアクターなので、これを精密に分析していきます。

・動物アクターの把握

人間の次に重要なアクターは動物です。

・アクター間の関係の把握

(3) 擬人化の認識

基本的に人物を特別扱いしたい。実際、話しをしたり、活動の中心は人物であります。だから、動物や物を擬人化して表現する事には特別なオブジェクト変換(+human という意味を付加したプロキシモデルを設ける)して処理していくことにします。

(4) メタファー関係認識

「中俣が来る」とか、「中俣ホームページ」とか、場所で人物や物事を把握することがよくあります。他にも、メタファーを用いた表現が沢山ありますが、そのメタファーが何を意味しているかを把握することです。

(5) 文脈把握

- ・ 欠落語の補充
- ・ 指示代名詞の補充
- ・ 同じ物を違う言葉で言っている事の把握

(6) カットの完成

・ 意味記号で知識が構成されているところを、具体的な単語で置き換え、カットとすること

(7) 推論の実現

- ・ リーズニング

「りんごを噛むと血が出る」なんていう文章が成立するかどうかは、「りんごを噛む」、「血が出る」という知識が存在するか、それらの間に因果律はあり得るかを知識と照合することになります。その照合がリーズニングです。

- ・ 試行錯誤で問題を解決すること

思考の大きな機能は、試行錯誤です。目標に至る経路を切り出しては、評価していく。基本的にはリーズニングと同じですが、意志の配下、プランニングしながら行っていくところが大きかりになっています。

- ・ 数学的推論

集合論処理、シミュレーションを実現します。

2. 日本語解析と生成に必要な知識

データ駆動型なシステムにする必要があるので、データを如何にきめ細かくもてるかが性能を決める大きなファクターになります。

(1) オブジェクト(オントロジー、モデル、カテゴリー)

アクター、属性、行為を表す各単語とそれに付属するフレームです。

(2) オブジェクト間の関係

連想全般を表します。

(3) 文法

単語、構文、係り受け情報、意味情報です。

(5) コーパス

表層文セット、深層文セット、文章パターンです。

これらは、知識オブジェクトのメソッドとしてインプリメントするとエレガントになります。

3.1 「発火」の分析

(1) ステージについて

「に出る」、「に入る」、「を持つ」----->has_a 関係知識を発火
「を出る」、「から去る」----->has_a 関係からの削除パターンの発火
現行ステージには弛緩法による最も発火したオブジェクトを選ぶ。

(2) シーン

名詞と一致するオブジェクトを発火します。メタファー関係の発火も行います。それでアクターを収集していきます。

(3) 擬人化

「板がご飯をたべた。」は「ご飯を食べる」というのが人間的な表現だから、「板」が擬人化されていると判断する。この辺も弛緩法で考えていく事になります。

(4) メタファー

発火による解決しかない。

(5) 文脈把握

アクター、ステージ、シーン、カットの発火関係の大きい物を選択して、指示代名詞解決、欠落語補填をしていく。まさに発火が解決策です。

(6) カットの完成

変数の埋め込みは文法を発火で選んでいくことで可能です。

(7) 推論

知識があるかどうかは「発火」によって検索していくしかない。数学的推論は発火によらないプロセスでの処理になります。でも、

4 . 分析 (その 1)

ステージの解析を説明します。

【ゲームの風景】

台にトランプを置く。

部屋 has (human, 台、 トランプ) Stage(部屋) Actor(human)

台 has(トランプ)

点棒を台に置く。

台 has(トランプ、 点棒)

部屋 has (human, 台、 トランプ、 点棒)

点棒をひっくり返して

部屋 has(human, 床、 台、 トランプ、 点棒)

床に落とした。

台 has (トランプ)

床 has(点棒、 human, 台)

点棒を拾って手に持った。

部屋 has(human, 床、 台、 トランプ、 点棒)

human has(手)、手 has(点棒)

・
・
・

という具合にステージ、シーンの解析をしていきます。自然言語処理もこれをやっていかななくてはなりません。

5 . 分析 (その 2)

発火の機構を説明します。

【列車の旅】

「 信子は牟礼駅まで歩いた。

牟礼で汽車に乗った。

明雄が駅に待っていて一緒になった。

二人は一緒に長野に行ったのだった。

長野では本屋とデパートに行った。

」

・「歩く」から、知識「(+human)-は-(+location)-まで-歩く」が発火する。

この変数に「信子」と「牟礼駅」を埋め込む。このとき、ステージ「家から牟礼駅」も発火する。ステージも日常の学習から知識化しているのである。

・アクター (信子、牟礼駅)

・「乗った」から、知識「(+human)-は-(+location)-で-(+object)-に-乗る」が発火する。

この変数に「信子」と「牟礼」、「汽車」を埋め込む。主格が欠落しているから、前にアクターの情報から「信子」を補填する。また、「牟礼」は「牟礼駅」とメタファー推論で相同と判断する。

- ・アクター（信子、牟礼駅、汽車）
- ・ステージ（牟礼駅）

「待つ」から、知識「(+human)-が-(+location)-に-待つ」が発火する。

「一緒になる」から、知識「(+human)-が-(+human)-と-一緒になる」が発火する。

この変数に「明雄」と「駅」、「信子」が埋め込まれる。「信子」は欠落語として、アクターから持ってきて補填される。「駅」はメタファー推論でアクターの「牟礼駅」と相同と判断される。

- ・アクター（信子、明雄、牟礼駅、汽車）
- ・ステージ（牟礼駅）

・「行く」から、知識「(+human)-は-(+location)-に-行く」が発火する。

この変数に「二人」と「長野」が埋め込まれる。「行く」の属性に「一緒に」が埋め込まれる。

「二人」は「明雄」と「信子」であることは、アクターの中味から推論する。

「長野」と「汽車」から「長野」は「長野駅」と相同であることを推論する。

- ・アクター（信子、明雄、牟礼駅、汽車、長野）
- ・ステージ（牟礼駅から長野駅）

・「行く」から、知識「(+human)-は-(+location)-では-(+location)-に-行く」が発火する。

・「と」から並置関係解析が発火する。

この変数に「信子」「明雄」と「長野」、「本屋」「デパート」が埋め込まれる。

アクターから「牟礼駅」と「汽車」が削除される。それは、長野市内の状況であるからである。ステージ解析の結果である（ステージ遷移の規則が知識にある）。

- ・アクター（信子、明雄、長野、本屋、デパート）
- ・ステージ（長野市街）

6．発火システムについて

ハードウェア的に XQuery/XPath を実現するのがビジネス的には有利であると思います。XML と XQuery/XPath は確立され、ビジネスの世界では良く使われているからです。

ロボットは、おかれた状況を常に解析していかなばなりません。状況は、知識オブジェクトの連想として全て表現できるでしょうか。

連想は基本的に2項関係ですが、配置など2次元マトリクス、3次元マトリクスになるかもしれません。

連想で全てが解決するならば、発火によって全ての知識処理の制御が実現できる事になります。

おわり