

考察：「動画システム」

2年前に論文「動画システム」を書き上げて、今やっとプログラミングに掛かれるまでになりました。当時は曖昧性管理を考えていませんでしたし、場所とか時間とかの包含関係も考慮していませんでしたが、今はそれを含めて挑戦している所です。

(例文) 昨日長野に行った。：ステージ(昨日、長野)を起こす。

レストランに入った。：ステージ(レストラン)を(昨日、長野)のサブとする。

食事をした。：ステージ(レストラン)内のカット。

レストランを出た。：ステージ(長野)に戻る。

ということ进行处理するのは、レストランが長野に包含されることを知らねばなりません。ステージの切れ目、シーンの切れ目は大抵なんとか分かるものですが、曖昧性はぬぐえませんね。

(例文) 柳原を発って長野を目指してあるいた。：何となく(柳原、長野)の中間がステージになりそうです。それには意味解析しないとなります。意味解析に曖昧性はつきものです。

それに、「長野」は「長野市」、「長野の街」、「長野駅」・・・「長野県」と、これまた曖昧です。文脈から判断しなくてはなりません。

今回こういったことを考えてみます。

第1章 意味の包含関係の解析

単語の意味の包含関係はどう構築していったら良いでしょうか。場所を表す単語で考えてみます。

長野の街はレストランとか、本屋さんとか、県庁とかで構成されています。レストランはテーブルとか人とかで構成されています。こんな風に考えていきますと、集合論で考えるのがよさそうと見えてきます。

釘とか木とか、板、さらにテーブルとか人とか、さらにさらにレストランとか本屋さん、県庁、長野の街、長野駅、日本、・・・と数え上げていきます。これらを「素材集合」ということにしましょう。

で、レストランを構成する集合を考えてみます。テーブルとか人、椅子、釘とか木、板、・・・と集合ができます。椅子は釘とか木、板なんかが集合となります。長野の街ですが、県庁、レストラン、椅子、釘、木、板、人・・・ですね。

この集合の包含関係で同格か、サブかスーパーかを判断できます。同格は排他要素がある場合、サブは全ての要素が含まれているもの、スーパーは自分の集合を全て含むものと定義します。同格は難しく、完全排他もあるわけですが、部分的に共有する集合要素があったりします。仙台の街は長野の街と殆ど共有するものが一緒でしょう。でも、場所座標が違います。ある視点では同じものですが、ある視点では排他の単語なのです。

包含関係は文脈依存ということなのです。

さらに考えるべきことは、計算機は有限の要素しか扱えません。プリミティブを「素材集合」の基盤にすべきです。ですが、プリミティブは必要に応じてブレークダウンしていくでしょう。つまり集合は学習されていくべきだということです。そんな動的なシステムとして、意味の包含関係は実現されるものです。つまり、

意味の包含関係 = (素材集合、視点)
と表現できます。

第2章 動画システムの曖昧性管理

動画システムはステージ（時間、場所）、シーン（アクター、フォーカス）、カット（1文の意味表現）の時系列です。基本的に時間と場所、アクターが重要な管理項目になります。時間は時系列として表現できますが、話が戻ったりします。そういうことに対応できなくてはなりません。場所は基本的に場所を表す名詞をトレースすることで解析しますが、包含関係が重要な解析項目になります。長野に話題があったとき、「仙台へ行った」ですと、長野からの移動で、ステージは同格なものへの推移ですが、「レストランへ行った」はサブなステージへの推移です。本ステージが長野であることに変わりはありません。それを追求しなくてはならないのです。

アクターですが、「本屋で本を買った」「本は安かった」という2つの文が重なっていたとき、前の文の「本」と後の文の「本」とは同じものという解析をしなくてはなりません。同様に、欠落語（主語、目的語がよく省略されます）の推定、指示代名詞の同定と同じような処理です。これも曖昧性を持って、管理していく必要があります。これ結構大きな処理です。

「本屋で白い本と黄色い本をかった」「白いのは安く。黄色いのは高かった」という場合の解析処理を考えてみてください。修飾語の同格さも解析しなくてはならないのです。しかもどんなに頑張っても曖昧性はぬぐえない。

この辺の議論は未夢プロジェクトが成功したときにお話していくとして、今回は大まかな曖昧性管理の枠組みを語りたと思います。

シーン、ステージ、カットを3つの時系列オブジェクトと捉えます。次の図の通りです。

ステージ																		
シーン																		
カット																		

----->時系列

図1 シーン・ステージ・カットの時系列

時系列の各升目には次のような曖昧性管理オブジェクトが入り、曖昧性管理オブジェク

トのエントリに推定の時間と場所、アクターが管理されます。

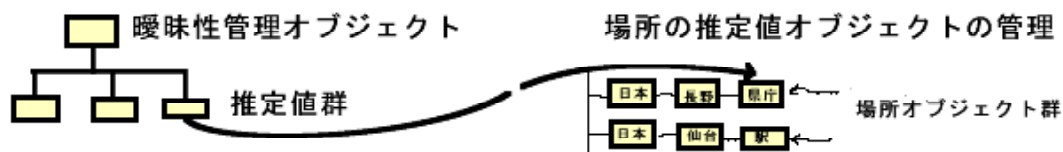


図 2 場所の管理

ま、とにかく大がかりに成ります。

第 3 章 曖昧性管理機構

曖昧性は何時か解消すべきものです。鍵と鍵穴システムで処理しようとしたとき、この曖昧オブジェクト群の内、これは消し込みましょう、これは推奨しましょうということが起きます。そのとき、解消した曖昧性がすみやかに他の関連曖昧性管理オブジェクトに反映でき、派生物も曖昧性除去できなくてはなりません。

ここで考えたのは、曖昧性オブジェクトをプールしていくことでした。実際の曖昧性のあるデータオブジェクトにはこの曖昧性管理オブジェクトを担わせるのです。派生オブジェクトは従いまして、派生元の曖昧オブジェクトと独自の曖昧オブジェクトの両方を持ちます。派生先であるいは派生もとで、項目削除となったときは、その旨を曖昧オブジェクトに記しますと、他の関連オブジェクト全てが削除対象にすることができます。簡単のために、曖昧オブジェクトをネストして管理しています。見かけ上一つの曖昧オブジェクトですが、相互に関連を自動的につけることができます。曖昧性管理はだから、シングルTONのファクトリーパターンを採用して実現しています。

おわり