

情景描写文からの対象世界の復元

山田 篤, 西田 豊明, 堂下 修司

京都大学工学部情報工学教室

本報告では空間的な情景を描写した自然言語文章を入力として, 対象間の関係を解釈して対象世界のモデルを構成する実験システム SPRINT について述べる. 文章中に現われる対象間の関係を論理的に一意な関係として表現するだけでは, 対象世界の状況は決して理解できない. 我々は, 対象世界に関する知識を解釈の過程に取り込むことによって, 能動的に解釈を行なうことを目指している. 空間状況記述には, 論理的に一意な解釈をすると矛盾が起こるが, 緩く解釈することによってもっともらしい対象空間の状況の解釈が可能となるような制約が多く含まれている. 我々はこれに対応するために, 論理的に一意な解釈からのずれとその解釈の尤もらしさの程度の相関を, 対象の空間的な属性に関する数値的な制約として表している. この制約を用いて, 対象世界全体の尤もらしさが最も大きくなるような解釈を導き出すことができる. 新しい制約が入ってくると, その解釈の尤もらしさを追加して全体としての尤もらしさが最大となるように, 解釈を漸次修正していく. 現在, 情景を記述した文章のサンプルを集めて, 評価を行なっている.

Interpretation of Spatial Descriptions Considering an Objective World

Atsushi Yamada, Toyooki Nishida and Shuji Doshita

Department of Information Science

Kyoto University
Sakyo-ku, Kyoto 606, Japan

This paper describes SPRINT, a computer program under construction which takes natural language texts in Japanese and produces a model of an objective world with 3 dimensions. We regard the world as an assembly of the objects, and represent the spatial concepts described in the text as numerical constraints on spatial attributes of objects in order to express the vagueness of spatial concepts. We use a theoretical device which we call the potential model for this sake. Interpreting the vagueness actively, we can get a concrete model of an objective world as an instance. The interpretation reflects the temporary belief about the objective world. Now the early version of SPRINT is running and images of an objective world can be obtained as an output.

1 はじめに

空間的概念は人間にとって非常に重要な概念である。人間は豊富な空間的概念をもち、これらは人間の抽象的な思考の一つの基盤になっていると考えられるが、人間がどのような空間的概念をもち、そのような概念をいかに取り扱っているかということについて、計算論的な見地からはまだ明らかにされていない。

本報告では具体的な空間に関する言語表現の理解について述べる。人間は言語表現（ある種の記号の表現）から記述された世界の状況の解釈としてのモデル（ある種のパターンの表現）を容易に構築することができる。我々はテキストの構造から対象世界の構造を再構成する過程に注目している。

我々は記述の解釈過程を、話者が意図した状況を能動的に再構成する過程とみなしている。実際に世界の具体的なモデルをつくることによって、対象世界に関する知識を解釈に反映させる。自然言語による記述は、記述された世界に関する様々な制約を含んでいると考えられる。これらの制約を対象世界の具体的な状況として解釈することによって、我々は対象世界に関する能動的な理解を目指す。我々の考えを実証するために、我々は実験システムSPRINT (SPatial Representation INTerpreter) をインプリメントした。

以下では、初めに空間的記述を解釈する際に生ずる問題点を指摘し、それらに対する我々のアプローチを示す。次に実験システムSPRINTの概観を示した後、具体例に即して説明する。

2 問題設定

自然言語の文章には、記述対象に関する様々な制約 (constraint) が表現されていると考えられる。自然言語では、対象世界の様々な具体的な状況が限られた語彙を用いて表現されるので、そこにあらわれている制約は、しばしば厳密でなかったり論理的でないことがある。

これらの制約について二つの問題がある。

一つは漠然性の問題である。ある種の制約は厳密に正確である必要はなく、それゆえに具体的な解釈では選択の幅をもっている。このような場合には、解釈と非解釈の間に明確な境界は存在しない。また、このような制約に対し論理的に一意な解釈を割り当てると、問題が生ずることがある。例えば、ある人が平安神宮と清水寺がともに京都駅の北東にあり、京都駅から平安神宮までは清水寺までの2倍の距離であることを知らされたとしよう。このとき、すべての制約を字義どおり厳密に解釈して、図1(a)に示すような対象世界の状況を思い浮かべ

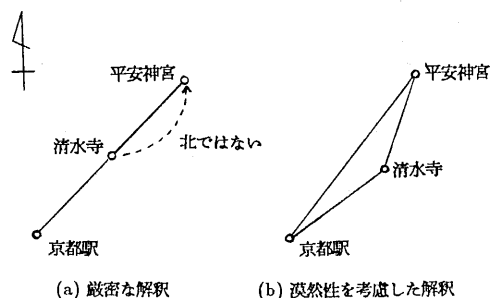


図1: 解釈における漠然性の積極的な利用

ることができる。今、「北東」を厳密に北東を意味すると解釈することに何の問題もない。次に、この人が新たに平安神宮は清水寺からみると北にあると教えられたとき、それはありえないと考えるだろうか。与えられた制約を厳密にとらえる限り、このような状況は確かに解釈できない。しかし、各制約を緩めて、「北東」は「厳密に北東」である必要はないと考えることによって、適当に解釈をずらして、例えば図1(b)に示すような状況を想像することは可能である。このような場合には、人間は漠然性を積極的に利用している。しかし、一つ一つの具体的な解釈の間には尤もらしさの違いがあり、ある解釈と他の解釈の尤もらしさを比べるようなある種の評価手段があると仮定してもよさそうである。また、どの制約を他の制約より確からしいと考えるかによっても、もっともらしい状況は異なってくる。

また、もう一つの問題として、文章中には明示的に与えられていない制約の存在がある。例えば、銀閣寺は京都大学の東にあるという制約を解釈しようとするとき、このままでは銀閣寺と京都大学間の距離に関する制約が欠如しているため、確からしいと考えられる適当な距離を考えなければ、具体的な状況の構築はできない。また、しばしば、我々は世界の具体的なモデルを構築するために、文章中には陽に書かれていないが、例えば、「建物が空中に浮いていることはなく、多くは地面に接して存在する」といった常識から導き出される制約を仮定しなければならない。

この他に開く[Hiraki88]は状況依存性の問題を指摘している。

3 アプローチ

我々は人間の振舞いを参考にしたいうえで、コンピュータモデルを用いて計算機上での空間的概念の取り扱いについて考える。

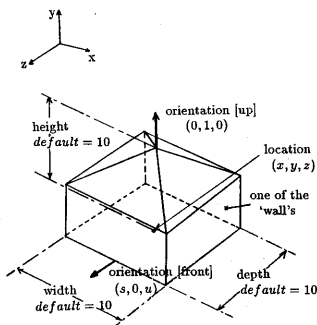


図 2: 「家」の空間的属性のパラメータ表示

通常の記事中の情景表現に明示的に表されている情報は、多くの場合わずかであるにもかかわらず、我々はかなり鮮明に情景を思い浮かべることができるように思われる。この一つの理由は、我々は相互に了解する暗黙的なプロトコルを利用しているためであろう。とともに、我々は我々自身の経験や感性に基づいて情報の暗黙的な部分を補うという創造的な作業を行なっていると考えられる。我々は以上のような考察のもとに、人間は常識・経験・主観などによって情景記述の暗黙的な部分を積極的に補うという仮説を立てている。

以下に示すように我々のアプローチはこの仮説を反映したものとなっている。

3.1 世界の幾何的表現

我々は対象世界をその構成要素としての対象の集まりとしてとらえ、各対象をその原型 (prototype) とパラメータ値の集合として表現する。具体的には各原型に対応するグラフィックオブジェクトを用意し、各グラフィックオブジェクトを、その詳細を決定するパラメータの値によって規定する。これによって我々の問題は、自然言語テキストからパラメータによって規定される対象の集まりへの変換の問題として定式化される。

例えば、「家」は図 2 のように家の原型と位置、向き、大きさなど、その原型を規定するパラメータの値によって表現する。

我々に与えられた問題は、与えられた文章を解析して、記述対象に対応するグラフィックオブジェクトを生成し、それらを規定するパラメータの値を決定する問題として特性化 (characterize) される。

3.2 空間に関する知識の表現のための機構

パラメータの値はコンピュータグラフィックスの場合と異なり、自然言語の文章が入力である場合には明示的に

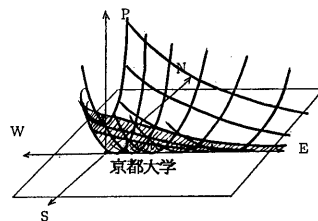


図 3: 「銀閣寺は京都大学の東にある」に対するポテンシャルエネルギー関数表示

与えられることはない。そこで文章中に表現されている対象に関する制約を、対象を規定するパラメータに関する制約として解釈し直す。

例えば、銀閣寺は京都大学の東にあるという制約を図 3 のように表す。ここで水平面は銀閣寺と京都大学の相対的な位置を表す 2 次元座標系、垂直面は京都大学に対する銀閣寺の相対的な位置がそこにあるときのものとの制約の受け入れ難さを表す。すなわち、銀閣寺の位置が京都大学の真東にあるときこの制約は最も受け入れ易く、方向がそこから遠ざかるにつれて容認しにくくなることが表されている。我々は、このようにある制約に関連するすべてのパラメータを引数としてとり、各解釈に対するコストを出力するような関数をポテンシャルエネルギー関数と呼び、このような表現をポテンシャルエネルギー関数表示と呼んでいる [YND88a,b]。これらは対象の幾何的なパラメータに関する数値的な制約である。ポテンシャル関数の値が小さいほど、解釈値としての幾何的なパラメータの組は確からしくなる。

ポテンシャルエネルギーという概念は元来、物理学のものである。一般に、ポテンシャルエネルギーを持つ力学系 (dynamical system) は非常に素直な性質を持つ。すなわち、自然界ではポテンシャルエネルギーの高いものは低い方へ変化しようとする傾向がある。その傾向の大きさはポテンシャル関数の勾配 (gradient) に比例する。本研究で用いているシミュレータは各地点で勾配を計算して、ポテンシャルエネルギーの低い方へパラメータ群の状態を変えるようになっている。このようにポテンシャルエネルギーという概念に基づくシステム構成をとっておくことによって、制約あるいは確からしさばかりでなく、現在の解釈が受け入れ難いとき、パラメータをどちらの方にどの程度変化させるとより確からしくなるかという手続的な情報まで簡潔に表すことができる。なお実際

のポテンシャルエネルギー関数は上のような位置パラメータばかりでなく、あらゆる種類のパラメータ空間上に設定することができる。

また、全ての制約をこのようにポテンシャルエネルギー関数として表すと一様性はあるが、元来漠然性を含まない制約にまでポテンシャルエネルギー関数表現を適用することは現実的でない。そこで論理的な制約も許すことにしている。

3.3 依存関係の手がかり

文章から得られる情報を対象間の空間的な関係を表す制約の単なる集まりと考えると、すべての制約を一度に考慮しなければならないという状況が生ずるが、これは条件が厳しすぎて、現実的ではない。

普段のコミュニケーションではそのような効率の悪いことはせずに、話者は聞き手が情景を再現するのを助ける情報を発話(utterance)に盛り込んでいるものと考えられる。

自然言語の文章を解析してみると確かにそのような情報は含まれている。例えば、上の銀閣寺と京都大学の例で、もしも京都大学がそれ以前の文脈にでてきており既知の対象となっていたならば、新たに現われた銀閣寺は京都大学に依存して考える。そこでこれをパラメータ間の依存関係(dependency)としてとらえ、パラメータ値の決定はこの順序に従って行なう。このように情景構築の手順指示を依存関係によって与える。

なお、テキスト中にあらわれる依存関係に関する手がかりは部分的なものであると考えられるので、実際には全体の依存関係を部分的な情報から決定する機構が必要になる。

また、パラメータ値の決定を依存関係の通りに一方向的に行なうと、あるパラメータ値が決定不能になることが有り得る。これを解決するためには依存関係を絶対的なものとは考えず、必要に応じて依存関係の順序とは逆に計算を行なうフィードバック手法を用いる。

3.4 依存関係に誘導された極小解釈の探索

ここに問題は情景を構成する対象の全パラメータの数の次元をもつパラメータ空間内でのポテンシャルエネルギー最小化問題に帰着される。パラメータ空間内でのパラメータのある組合せは現実を起こり得ない故に禁止されているかもしれない。ポテンシャル関数無制限に選んだ場合、ポテンシャルエネルギーを最小にするパラメータの組を求めることは一般に不可能である。そのため我々は現在、最急降下(gradients descent)法に基づく数

値的な極小探索アルゴリズムを用いている。

各パラメータの初期値がうまく与えられたならば、本手法はうまく動作する。我々はこの過程を求解というよりも、むしろ解の調整過程としてとらえている。

ここで我々が最も興味があるのはポテンシャルエネルギー関数全体の定性的な性質である。極小解は局所的にみたとき、その周辺において最も確からしい解釈に相当する。与えられた問題に対して極小解が一つなのか、それとも複数個あるのか、特に複数個ある場合には、それらの確からしさは同程度なのか、それともいくつかが際立って確からしさが高いのか、そのようなことを知りたいのである。

4 SPRINTによる処理の概観

—自然言語文章入力からのモデルの構築過程—

SPRINTは空間的な状況を記述した自然言語テキストを解釈して、対象世界の3次元モデルを構築する実験システムである。これは現在、人工知能用ワークステーションSymbolics 3650上にインプリメント中である(描画にはSymbolics社のS-Geometryを使用)。

SPRINTの中心は、自然言語テキストから抽出された記号的な制約から尤もらしさを考慮した対象世界の解釈を導き出すためのポテンシャルエネルギー関数である。

SPRINTの概要を図4に示す。処理の前半は自然言語テキストから論理的な制約を抽出する過程である。これはさらに言語解析と空間関係の抽出の2つのサブシステムからなる。

後半は対象間の制約を数値的に解釈し、対象世界の具体的な解釈を作成する過程である。これはポテンシャル関数作成とポテンシャル極小化、さらに描画の3つのサブシステムからなる。

4.1 統語解析

入力は日本語で記述された情景描写文とする。入力された文章を、統語解析サブシステムでパースして、表層格構造を得る。この段階で使う知識は、語の範疇に関する一般的な知識のみとする。

一般にある文に対する格構造には多くの可能性がある。統語的、辞書的な知識からはどの構造が正しいかを確定できないときには、SPRINTは可能な解釈をすべて生成し、以後の過程に処理を委ねる。

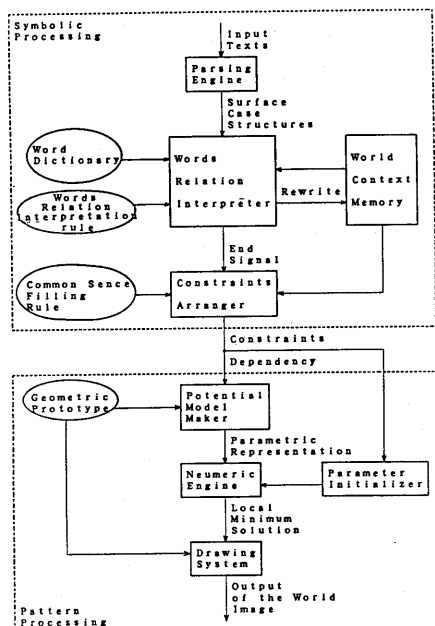


図4: SPRINTの概要

4.2 空間的関係の抽出

次に空間的関係の抽出サブシステムで語彙に関する空間的な知識を用いて、格構造から対象に関する空間的な制約と依存関係を得る。これはテキストに対する空間的な観点からの意味解析に相当する。さらに、文章レベルの談話解析を行い、語句の指示対象を特定する。同種の対象を表す語句が複数出現しているときに、それらが同一の対象を指しているのか否かを決定する。指示語の解析においては、既に生成した対象オブジェクトと対応をとる。対象の候補が複数個存在する場合には、オブジェクトを特定しなければならない。現在は、矛盾が生じない限り、統語構造において最も近くに現われた対象と対応をとり、必要に応じて選び直すという方針を用いている。

SPRINTは格構造において関連づけられた語を末端の語から解析することにより、対象間の制約を抽出する。格構造中のすべての枝が、「辞書」と「関係解釈ルール」を用いて解釈されていく。一時的な結果が「文脈記憶」に書き込まれていく。これは対象に関する知識とそれまでの「文脈記憶」の内容に応じて書き換えられていく。

SPRINTは格構造中のすべての関係が解釈されるまでこの過程を繰り返す。また、SPRINTは情報参照の順序を依存関係として抽出する。これらの解析はすべて

の可能な格構造に対して適用される。この段階で構文的に可能な格構造の内のいくつかは意味的な矛盾から取り除かれる。

空間的関係の抽出の際には、「上」「左」「東」等の空間的な関係を明示的に表す語が大きな手がかりとなる。我々はこのような語を空間語と呼んで対象を表す通常の名詞と区別している。空間語の多くは、「テーブルの上に」「その建物の南側に」等のように

普通名詞+格助詞「の」+空間語

の形であらわれる。このときの空間的な状況は、空間語のみによって規定されるのではなく、より詳しくは前後の語句によって定まることが多い。例えば、

「テーブルの上に灰皿がのっている」
 「テーブルの上にテーブルクロスを敷く」
 「テーブルの上に風船が浮かんでいる」

では、それぞれ空間語によって関係付けられている対象自身の性質や述語によって、空間語「上」のより具体的な解釈が決まる。我々は、空間語に接続する語句の範囲と対象間の関係との対応をルールの形で用意しておくことにより、空間的情景の構築を行う際に必要となる制約を作成している。

また、「ある」「立つ」「見える」等の動詞も空間的な関係を記述する。このうち「見える」のように、人間が視覚を通じて空間的な関係を把握するものに関しては、視点と視線に関する制約を抽出する。

4.3 空間的関係の表現

ポテンシャル関数作成サブシステムでは、初めに対象の幾何的表現の組を生成する。このために、対象をいくつかのパラメータスロットをもった幾何的オブジェクトに対応づける原型の集まりを予め用意している。

次に空間的関係の抽出サブシステムで得られた制約を対象に対応づけられた原型のパラメータに関する数値的な制約として解釈していく。我々のコンピュータモデルで表現可能なパラメータに関する制約は、漠然性をもつ制約、確定的な制約、禁止的な制約の3種である。

4.3.1 漠然性をもつ制約

漠然性をもつ制約をポテンシャル関数によって表す。このグループの制約はすべて、あるパラメータ群から計算される量が、ある特定の値のときに最も確からしく、その値から離れるに従い、確からしさが減少するものである。従って、これらの制約はパラメータ群からポテン

シャルを算出するためのポテンシャル関数によって表される。以下では、パラメータ群から計算される量がスカラー量であるか、ベクトル量であるかによって、二つの場合にわけて考える。

(a) スカラー量に対するポテンシャル関数

パラメータ群 $p_1, p_2, \dots, p_n$ から計算されるスカラー量 $l = f(p_1, p_2, \dots, p_n)$ が、値 v をもつことが最も確からしいとき、ポテンシャルは

$$Potential(p_1, p_2, \dots, p_n) = P(l) = \frac{K}{2} \left(\frac{l-v}{s} \right)^2$$

where, $K(> 0)$ is a gradient constant.

によって計算される。ここに、 s は必要な精度の違いを吸収するためのスケールファクタである。

(b) ベクトル量に関するポテンシャル関数

パラメータ群 $p_1, p_2, \dots, p_n$ から形成されるベクトル量 u がベクトル量 v に一致することが最も確からしいとき、 u と v のなす角は

$$\theta = \arccos \frac{u \cdot v}{|u| \cdot |v|}$$

であり、 θ が小さいほど、 u のとる値は確からしいと考えられる（我々は、ベクトル量に関する制約を向きを表すためにのみ用いており、ベクトルの大きさについては考慮していない）。よって、ポテンシャルは

$$Potential(p_1, p_2, \dots, p_n) = P(u) = \frac{K}{2} \arccos^2 \frac{u \cdot v}{|u| \cdot |v|}$$

where, $K(> 0)$ is a gradient constant.

となる。

いずれの場合にも、勾配係数 K の値を調整することで、その制約の確からしさに関する条件を取り入れることができる。

4.3.2 確定的な制約

偶然性をもたない確定的な制約を固定関数によって表す。このグループの制約は、あるパラメータが何らかの形で計算される特定の値をもたなければならないというもので、そのパラメータがもつべき値の計算式を与えることによって表す。この制約によって一意に定まったパラメータの値が、他の制約によって異なる値を要求されたときには、整合性が保たれるようにフィードバックを試みる。

4.3.3 禁止的な制約

パラメータの定義域の中で禁止される範囲を禁止領域として表す。

あるパラメータについて、取り得ない範囲を禁止領域として宣言する。禁止領域は禁止半開空間（境界面上の1点の座標と禁止領域を向く法線ベクトルによって特徴化）の組合せ（union と intersection）で表現する。

ポテンシャル極小化の過程で、ここからの脱出と再入阻止をはかる。

4.4 依存関係の作成

協調的なコミュニケーションでは情報の参照手順が文章の言語構造の中に情報の新旧という形で埋め込まれていると考えられる。そこで、言語解析時に保持しておいた情報の新旧に関する記録をもとに、まず対象間の参照関係を定める。次に、既に抽出された空間的關係から、ある対象間で実際に参照されているパラメータを選び、それらの間に参照側を子、被参照側を親とする依存関係のリンクを張る。

4.5 ポテンシャルを極小とする解の探索

最後にポテンシャル関数の極小解計算サブシステムがこれらの制約を満たしポテンシャルを極小とするようなパラメータの値の組を数値的に求める。

パラメータの値の計算は依存関係の親側から子側へと行なう。

親のパラメータがすべて確定したノードについて、パラメータの値を仮決めし、ポテンシャルエネルギーを低くする方向に各パラメータの値を変更していく。パラメータの値をどの様に変更してもポテンシャルがそれ以上低くならなくなれば、そこがそのパラメータのポテンシャル極小点であると考えられるので、そのときの値をパラメータの値として確定する。これを依存関係に従って各ノードのパラメータについて実行する。

依存関係に従った解釈の結果、局所的に確からしさの低い（ポテンシャルが不当に高い）解釈が出現した場合や、矛盾が発生した場合に、依存関係を逆にたどって、子側から親側へ影響を及ぼすフィードバック処理を施す。

フィードバック過程では親側のノードが子側に及ぼしているのと大きさは同じで向きが逆の力を子側から親側に作用させる。

最後にシステムは最終的な解釈結果を3次元ソリッドモデルを用いてグラフィックディスプレイ上に表示する。

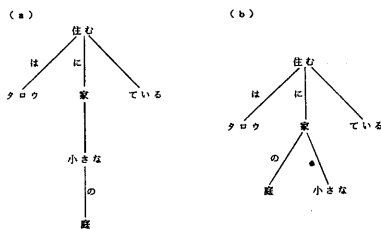


図5: 文(1)に対する表層格構造

5 詳細例

(case-1)

特に空間的な状況を記述した文ではないが、次の文について考えてみる。

(1) 「タロウは庭の小さな家に住んでいる。」

この文に対する表層格構造は図5のようになる。この文に対する表層格構造は二つあり、実際にそのいずれの解釈が正しいかは、構文的な知識からだけでは決めることができない。しかし、これは次の段階で、例えば、タロウが何者であるかを現在の文脈記憶から探してくることによって決定できる。即ち、タロウがそれまでにでてきた誰か人間の名前であるということがわかれば、とるべき格構造としては(a)の方が確からしくなり、タロウは自分の家の庭として小さな庭を持っているような家に住んでいると解釈することができる。また、タロウが例えば犬の名前であったならば、格構造としては(b)のほうが優先され、タロウはある庭の中に存在する小さな家に住んでいると解釈することができる。この場合、当然、どちらの解釈をとるかによって、この後抽出される制約は異なり、異なった空間的な状況が想定される。

(case-2)

SPRINTへの入力として次の文を考える。

(2-1) 「崖の縁に家がたっている。」

文(2-1)に対してパーサが格構造を生成した後(この文に関して構文解析レベルでの曖昧性はない)、語の接続を順次解釈して、文脈記憶にオブジェクト「家」がオブジェクト「崖」の部分である「縁」の位置に存在しているという情報を書き込む。次にSPRINTは対象世界に関するより具体的な解釈を行なう。まず、「崖」の原型が

OBJECT-1として用意される。システムは崖とその縁との間の関係を知っているので、OBJECT-1の「縁」がPART-1としてマークされる。次に、「家」の原型がOBJECT-2として用意される。OBJECT-2に関する制約は次のようになる:

- (a) OBJECT-2の底はOBJECT-1の上面に接している。
- (b) OBJECT-2の底はOBJECT-1の表面からはみでない。
- (c) OBJECT-2の位置はPART-1のなるべく近くに

ポテンシャルモデル生成部はこれらの制約を対象を規定するパラメータ間の数値的な制約として解釈する。制約(a)はOBJECT-1の位置と高さ、OBJECT-2の位置の間の固定関数として表現される。制約(b)はOBJECT-1の位置と大きさ、OBJECT-2の位置に関する禁止領域として表現される。制約(c)はOBJECT-2とPART-1の間の距離に関するポテンシャル関数として表現される。これらの様子を図6に示す。各パラメータに適当な初期値を与えて極小解を計算した結果をグラフィックスで表示すると図7のようになる。

次に以下の文を追加入力する。

(2-2) 「その家の南の窓から見おろすと、海が見える。」

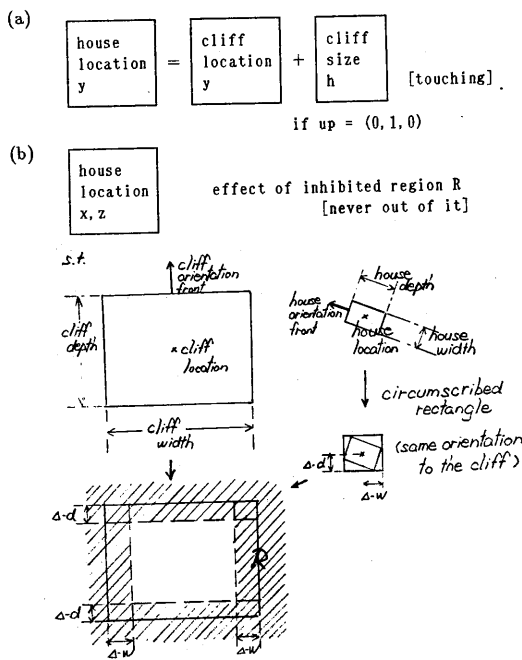
今回は「家の南の窓」の部分の表層格構造に曖昧性が現われる(図8)。しかし、これは次の段階で「家」と「窓」の単語辞書から両者の関係を考えることによって(b)が消去される。逆に、例えば、「窓」の代わりに「家の南の道」のような記述ならば、「家」と「道」の関係から(a)の構造が消されることになる。

以下、先と同様の処理を施して、最終的に図9に示すような出力を得る。

(case-3)

本報告の初めに上げた例を実際に取り上げてみる。

- (3-1) 「清水寺、平安神宮は京都駅からみると、北東の方角にある。」
- (3-2) 「京都駅から平安神宮までは清水寺までの2倍の距離がある。」
- (3-3) 「平安神宮は清水寺からは南の方角にあたる。」



(c) house location (x, y, z)
 edge endpoints $(x_1, y_1, z_1), (x_2, y_2, z_2)$
 These are calculated from the edge location, size and orientation.

$$\text{line } \frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}$$

distance between house and edge

$$d = \sqrt{\frac{(x_2 - x_1)(y_2 - y_1 - z_1) - (x_2 - x_1)(y_1 - z_1) + ((x_2 - x_1)(y_2 - y_1) - (x_2 - x_1)(y_1 - z_1))^2 + ((x_2 - x_1)(y_2 - y_1) - (x_2 - x_1)(y_1 - z_1))^2 - (x_2 - x_1)(y_2 - y_1 - z_1)^2}{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}}$$

The nearer, the better.
 potential function

$$P(x, y, z) = \frac{K}{2} d^2$$

図6: 文(2-1) に対する制約の表現

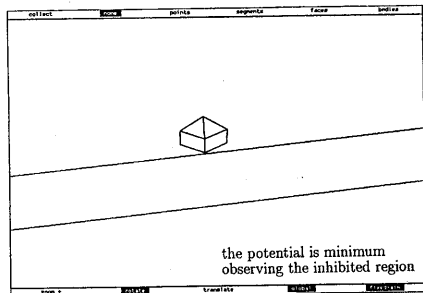


図7: 文(2-1) の解釈結果

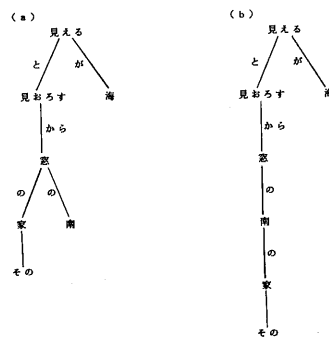


図8: 文(2-2) に対する格構造

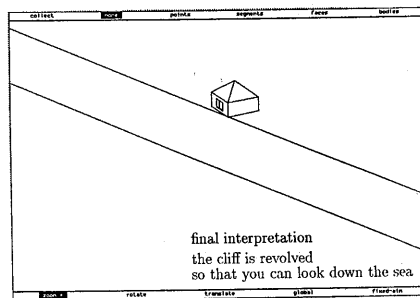


図9: 文(2-2) の解釈結果

これらは文としては比較的簡単なため、言語的な解析の部分の説明は省略する。

文(3-1) と(3-2) が入力された時点での対象世界の解釈は図10のようになる。このとき、それぞれの制約はほぼ厳密に満たされている。

次に文(3-3)を追加する。このままでは、もっともらしい解釈ができないので、システムは対象世界のモデルに修正を加えて、図11のような結果を得る。

このとき、更に各制約の勾配係数をかえてみる。まず、北東という制約を二つとも緩めてみると図12(a),

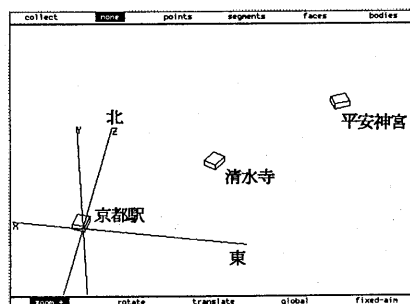


図10: 文(3-1), (3-2) に対する解釈結果

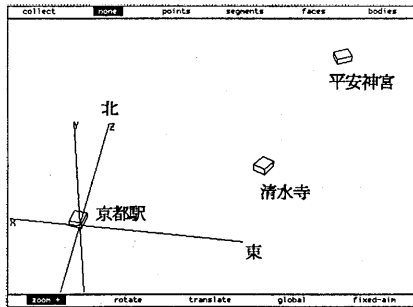


図 11: 文(3-3) まで入力したときの解釈結果

平安神宮の方だけを緩めてみると図12(b) のようになる。更に距離が2 倍という制約のみを緩めると図 12(c), 北という制約を緩めると図 12(d) のようになる。

6 検討

これまでに空間概念に関して言語的な側面からのアプローチとしては[Okada73,76,86] [Hersk86] などがある。一方、コンピュータモデルの構築に関しては、典型的な論理構造をもつだけでは不十分である。もっとも典型的なものは[Novak77] や[Waltz81] にみられるようなスロットフィラー表現であるが、実際には漠然性を含んだ記述から論理的に一意な構造を抽出することは困難である。

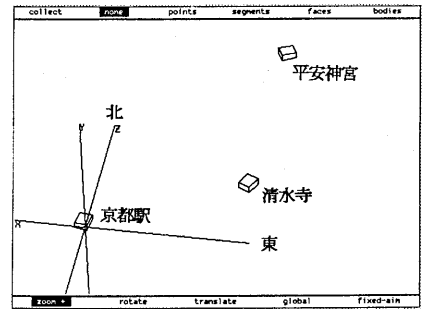
[Davis81] では *fuzz box* を用いることによって漠然性の問題を取り扱っているが、これは基本的に領域の考えに基づいており、解釈の確からしさの程度は考えられず、我々の問題設定である能動的な文章理解にはそれだけでは使えない。

我々のアプローチでは意図された状況に応じて、連続的、または不連続的に変化する尤もらしさを表すための数値的な計算モデルを用いて、論理的な制約と漠然性をもつ制約の統合をはかっている。処理の過程で主観が入ることは否めないが、ある種の主観無しにはこのようなタスクの実行はむずかしいと思われる。

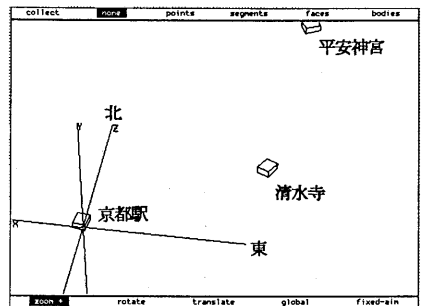
現在、情景を記述した文章のサンプルを集め評価を行なっている。これをもとにして、自然言語の空間的な解釈に必要な辞書の構築と、空間的制約のポテンシャル関数表示を実験的に行なっていく。

また、今後の課題としては、以下の点があげられる。

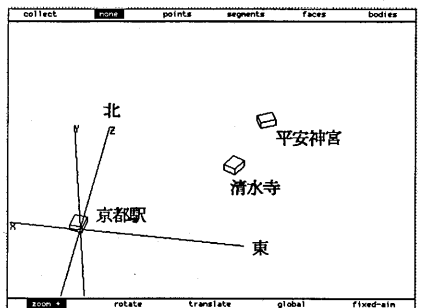
- (i) 作成したモデルの解釈過程での積極的な利用
曖昧な文章に対し複数の解釈の可能性があるときに、解釈の過程でそのときの文脈記憶の内容や実際に作成したモデルなどから、確からしさの順位が変わ



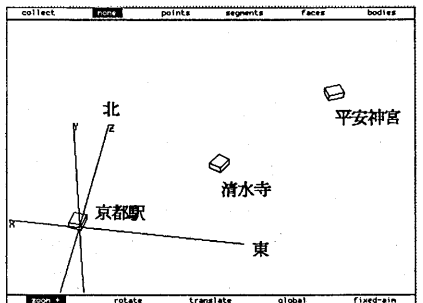
(a) 「北東」という制約を緩めた結果



(b) 平安神宮のみ「北東」という制約を緩めた結果



(c) 「2 倍」という制約を緩めた結果



(d) 「北」という制約を緩めた結果

図 12: 各制約の確信度の変更

る場合、統合的に処理をすすめることにより、それらの間での最尤解釈の移行をスムーズに行う。

(ii) 一般的な概念の空間的アナロジーの分析

より一般的な抽象概念の取り扱いにおいて現われる空間的な概念を用いたアナログ的な言語使用を、具体的な空間概念をもとに分析する。

7 おわりに

本報告では、自然言語テキストを入力として、その解釈を行ない対象世界の3次元モデルを再構成する実験システムの概要について述べた。本研究の手法をまとめると、以下ようになる。

1. 情景はパラメータ化した対象 (object) の集まりとして記述する。
2. 自然言語表現の意味を対象のパラメータの組に関する数値的な制約 (constraint) として表現する。
3. 漠然性を持つ制約はポテンシャルエネルギー関数によって表す。
4. 制約を利用して情景を再現していく手順に関する情報は自然言語表現の中に含まれていると考え、これを抽出して情景再現に利用する。
5. 情景はその場その場で最適と考えられるものを逐次組み立てていく。暫定的な情景を後の文の内容と照合して漸次修正していく。

現在、モデルとして表現しているのは、対象の幾何的な属性のみであるが、他の物理的、機能的属性に対しても同様のアプローチが有効であると考えられる。

参考文献

- [Davis81] E.Davis. Organizing Spatial Knowledge. Yale University, 1981.
- [Hersk86] A.Herskovits. Language and Spatial Cognition. Cambridge University Press, 1986.
- [Hiraki88] 開, 安西. 空間的関係の表現. 日本ソフトウェア科学会大会論文集. 1988.
- [Novak77] G.S.Novak Jr. Representations of knowledge in a program for solving physics problems. In *Proceedings IJCAI-77*.
- [Okada73] 岡田, 田町. 自然語および図形解釈のための単純事象概念の分析及び分類. 電子通信学会論文誌 Vol.56-D, No.9, pp.523-530, 1973.
- [Okada76] 岡田, 田町. 図形の意味解釈とその自然語

記述—要素的図形認識と構造分析. 電子通信学会論文誌 Vol.59-D, No.5, pp.323-330, 1976.

[Okada86] N.Okada. Towards a unified understanding of natural language and picture patterns. In *Proc. Intl. Sympo. on Language and Artificial Intelligence, Kyoto March 16-21, 1986*.

[Waltz81] D.L.Waltz. Towards a detailed model of processing for language describing the physical world. In *Proceedings IJCAI-81*.

[YND88a] 山田, 西田, 堂下. 漠然性を含む空間状況解釈のためのポテンシャル極小化アプローチ. 情報処理学会知識工学と人工知能研究会, 58-3, 1988.

[YND88b] 山田, 西田, 堂下. 2次元平面におけるポテンシャルモデルを用いた位置関係の推定. 情報処理学会論文誌, Vol.29, No.9, pp.824-834, 1988.