

エージェントのための日本語における空間表現の定式化に関する研究

小山 智史[†] 齋藤 豪^{††} 中嶋 正之[†]

エージェントに自然言語による指示を理解させるためには、自然言語が持つ漠然性、曖昧性の処理を行わなければならない。本稿では指示文中における空間表現に着目する。自然言語による空間表現は漠然とした領域を示す。一方、エージェントが動作を行う時にはある一点を必要とする。そこで各空間表現に対応したポテンシャルモデルを導入することにより、空間表現を広がりを持ちかつ数値的に扱うことを可能にした。これを Space オブジェクトとしてエージェントシステムに実装し、その有用性を確かめた。

Modeling Method of Japanese Spatial Prepositions for Autonomous Agent

TOMOFUMI KOYAMA[†] SUGURU SAITO^{††} and MASAYUKI NAKAJIMA[†]

When communicating with animated agents through natural language dialogue, it is necessary to deal with vagueness and ambiguity. This paper focus on the vagueness of spatial prepositions and proposes a representation of a location. Though spatial expressions in natural language represent vague space, agents need concrete coordinates to do the order. So we use potential model in order to bridge the gap between vague space and concrete coordinates. Through the implementation of a prototype system, we show our model has utility.

1. はじめに

近年自然言語を理解し行動するバーチャルエージェントについての研究が盛んに行われており、自然言語で指示を与えられるエージェントはヒューマンインターフェースとして非常に有用であると考えられている^{1)~3)}。エージェントが自然言語による指示を理解し実行するには、空間を表現する言葉を扱うことは避けられない。しかし、空間を表現する言葉は概念を記号化したもので漠然性と曖昧性が含まれており、その解消が大きな問題となる。

例えば次の対話を考えてみる。

Human: 机の前にボールはある?

Agent: あります。

Human: それを右に置いて。

まず、「机の前」が表す空間にはいくつかの候補がある。Human から見て前、Agent から見て前、机自体が方向を持っているときの机自体の前、の3種類である。次に「それ」が何を指しているのかを同定しなければならない。この対話で出てきた名詞は机とボールである。これらの作業は曖昧性の解消を行うということである。「置く場所」について「右」から決定することが必要であるが、「右」が表す空間にはある広がりがある漠然性を持つ。この広がりの中から「置くための一点」を決定することが漠然性の解消である。

このように簡単な対話であっても多くの漠然性、曖昧

性を含んでいることがわかる。本稿ではエージェントが行動するために、空間表現の漠然性を解消するための仕組みを提案する。

次節では自然言語とエージェントについて行われている研究を紹介する。次に空間表現についてのまとめを行い、実装の仕組みを紹介する。最後に実際のバーチャルエージェントシステムでの動作結果を示し、考察を行う。

2. 関連研究

2.1 自然言語からの空間構築

空間状況を自然言語で表したテキストから仮想空間を構築するシステムとして山田らのSPRINT^{4)~6)}がある。SPRINT ではポテンシャルモデルを用いて空間表現の漠然性を再現し、対象空間を数値的に扱っている。ポテンシャルモデルとはある対象物体のその位置での存在を仮定した場合の過程の受け入れにくさを表現しており、ポテンシャルが小さいほどその存在が確からしいことを表す。本稿ではこのポテンシャルモデルを参考に拡張性の高いポテンシャルモデルを考えることで、空間表現にとどまらず、エージェントの動作における漠然性も統一的に扱うことを可能にする。

2.2 空間状況の自然言語による表現

仮想空間の状況を自然言語で表現するための研究として八嶋らの研究^{8),9)}がある。この研究ではエージェントが仮想空間内の物体の位置を、他の物体を用いて表現することを目的としている。そのため、八嶋らは空間表現をSPRINT で用いられたポテンシャル場によって定式化を行い、空間の漠然性を数値的に扱った。これによりエージェントは「机の前」という漠然性を持った空間を理解することができるようになり、「机の前にボールはある?」という命令には対応が可能となった。しかし、漠

[†] 東京工業大学 大学院 情報理工学研究所
Graduate school of Information Science & Engineering,
Tokyo Institute of Technology

^{††} 東京工業大学 精密工学研究所
Precision and Intelligence Laboratory, Tokyo Institute of
Technology

然性を数値的に表現したのみであり、エージェントが動作を行うために必要な、漠然性の解消を可能にするまでは至っていない。本稿では漠然性の解消も可能とするポテンシャル場を定義する。これにより漠然性を持つ「机の前」の空間から状況に応じて一点を選び出すことができるようになり、「机の前にボールを置いて」といったような命令にも対応が可能になる。

2.3 知的エージェントシステムに関する研究

日本語による操作を可能としたエージェントシステムとして傀儡⁷⁾が開発されている。傀儡システムでユーザに与えられる課題は、ユーザが仮想世界の映像を見ながらその世界内に存在するソフトウェアロボットを使って仮想世界内の物体を正しい位置に配置することである。傀儡ではエージェントの動作プランを記号表現で決定することに重点が置かれており空間概念については深く取り扱っていない。そこで本稿では空間概念について定式化し、漠然性と曖昧性を解消することで、記号表現で考えられた動作プランから実際のエージェントのアニメーションを出力するための補助を行う。

3. Space

3.1 エージェントシステムについて

まず自然言語による指示が可能なバーチャルエージェントシステムのアーキテクチャについて説明を行い、本稿で提案する手法の位置付けを明らかにする。

今回我々が使用するエージェントシステムの概要を図1に示す。音声により指示が入力されると、音声認識、構文解析、発話意図解釈が行われ、エージェントは指示の目的を満たすための動作を決定する。この動作を決定する機能をマクロプランニングと呼ぶ。マクロプランニングまでは記号の論理演算のみでの処理が可能である。一方、この動作を実際に目に見える形で実行する際には記号表現ではなく、仮想空間内の具体的な座標によってエージェントの動作が定義される。そこで記号を具体的な座標に読み替える必要が発生する。本稿は空間表現を定式化し数値的に取り扱うことで記号を具体的な座標に変換し、空間表現を含む記号の動作列からエージェントのアニメーションを出力する事を目的とする。これにより、エージェントシステムの研究は記号表現によるプランニングと、座標からのアニメーション生成という二つの部分に分けて進めることが可能になる。

3.2 日本語による空間表現

まず本稿で取り扱う空間表現について述べる。日本語による空間表現は次の2通りに分類される¹⁰⁾。

- 空間名詞+空間辞
「前に」「中で」など
- 空間辞
「～に」「～を」など

「空間辞」により空間を表現する場合には、空間辞の前後の名詞、動詞に関する知識が必要となる。例えば「椅子に乘せる」という表現では普通「椅子の座る部分の上に置く」という解釈が行われる。これは椅子の形状に関する知識がなければ不可能な解釈である。そのため本稿では「空間名詞+空間辞」の形を取り扱う。空間辞の種類やあるいは動詞により表現される空間が異なることがあるが、

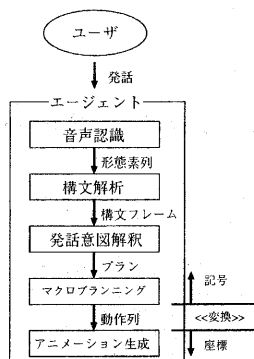


図1 エージェントシステムの概要

空間名詞には基本となる概念が存在し、そこからいくらかの例外を許容することで様々な状態を我々は理解することができる¹¹⁾。そこでまず空間名詞ごとの基本的な概念を定式化する。本稿では「空間名詞+空間辞」に「+状態動詞ある」を基本となる空間表現とし、この語句が表す概念を空間名詞の基本概念として取り扱う。

空間名詞には方向の制約を表すもの、距離の制約を表すもの、位置の制約を表すものの3つに分類できる。以下にその例を示す。

- 方向の制約を表すもの
前、後、左、右、上、下、向こう、手前など
- 距離の制約を表すもの
近く、そば、遠く、付近、麓、隣など
- 位置の制約を表すもの
内、外、中、まわり、間、頂上、底など

3.3 Space オブジェクト

空間表現をエージェントが処理する際に最も問題となるのは空間名詞が表現する空間が漠然性、曖昧性を持つ点である。同じ「机の前」という表現であっても、空間の状況によって指し示されている空間は異なる。また、「机の前」とは一点ではなく、ある程度の連続的な広がりをもった領域を表現するため、あらかじめある一点を「机の前」として決めておくことはできない。

本稿では、エージェントが指示文中に現れる空間表現を処理するためのオブジェクトとして Space オブジェクトを提案する。例えば「机の前まで歩け」という指示では、エージェントは「机の前」という空間表現を「前」という空間名詞に対応した Space オブジェクトを取り扱うことで理解する。また、空間名詞の他に指示文から暗に空間的制約を受けることがある。「机の上のボールを取って」という指示では前提として「ボールに手が届く」空間に移動することが必要である。この「手が届く」空間のようにエージェントが行動する際に必要な空間的制約も空間名詞と同様に Space オブジェクトとして表現する。Space オブジェクトによりエージェントは空間表現を理解し、指示された行動を自律的に生成することが可能となる。

Space オブジェクトに求められる機能は次のものが挙げられる。

- (1) ある物体が、Space オブジェクトで表現される空

間内にあるかどうか判定する

- (2) Space オブジェクトにより制約された空間内で行動を行うための一点を決定する
- (3) Space オブジェクト同士の AND や OR や NOT をとることができる

これらの機能を実現するような Space を設計する。

3.4 ポテンシャル関数

Space オブジェクトでは空間表現の持つ漠然性、連続性をポテンシャル関数で表現する。ポテンシャル関数とは空間の各座標において、どの程度空間名詞の表現する概念に相当であるかを示す値を決める関数である。ポテンシャル関数は次の条件を満たすように設計する。

- (1) ポテンシャル関数は空間内の任意の点で微分可能である
- (2) ポテンシャル関数は0~1の値をとり、値が1の時間がSpaceオブジェクトの表現する概念に最も適当である

「机の前まで歩け」という命令がエージェントに与えられた時、アニメーション生成部では「机の前」のある一点を動作の終点として必要とする。この一点はポテンシャル最大の地点であるとと考えられ、このポテンシャル最大の点を最急降下法を用いて計算を行う。最急降下法を用いるため、ポテンシャル関数は任意の点で微分可能でなければならない。

「机の前で机の上のボールを取って」という指示では、エージェントは「机の前」にいて、かつ「ボールに手が届く」ことが求められる。このように2つあるいはそれ以上の空間的制約が同時に存在するとき、このポテンシャルはそれぞれの空間的制約に対応したSpaceを合成することで指示文中の空間的制約に対応したSpaceオブジェクトを生成する。新しく作られたSpaceオブジェクトのポテンシャル関数はもとのSpaceのポテンシャル関数の積で構成される。ポテンシャル関数の取る値を0~1とし、ポテンシャルが大きいほどSpaceが表す概念に適すると定義することで複数のSpaceの合成により作られるポテンシャル関数でももとの全てのSpaceでのポテンシャルが大きい地点のみポテンシャルが大きくなる。

ある物体が空間表現により制約された空間内にあるかどうかの判定では、空間表現に対応したSpaceのポテンシャル関数を用いて物体が存在する位置のポテンシャルを計算し、その値で判定することが可能である。

3.5 Space が持つインターフェース

以上の性質を満たすSpaceを生成するためにSpaceオブジェクトは以下のインターフェースを備える。

- (1) potential(coordinate)
座標 coordinate 上でのポテンシャルを計算する
- (2) optimize()
最もポテンシャルが高い座標を探して返す
- (3) getConjunctionwith(Space)
他のSpaceオブジェクトとの合成を行ったSpaceを生成する
- (4) getNotof()
自分自身のNotを表すSpaceを生成する

4. Space オブジェクトを実装したエージェントシステム

本節では各空間表現に対応したSpaceの定義とエージェントシステムへの実装方法について説明する。エージェントシステムの概要は図1である。

Space オブジェクトはエージェントの外界認知機能として、マクロプランニング部で生成される。例えば、「机の上に青いボールはあるか?」という問いに対して、プランニング部は「机の上」というSpaceオブジェクトを作り、ボールのうち「机の上」Space内にあるものが存在するかどうかをチェックする。このときマクロプランニング部ではボールや机の具体的な座標を扱う必要が無いことが重要である。

「机の前まで歩け」という命令に対しては「机の前」に「歩く」という動作をプランニングする。実際の動作を出力するのはアニメーション生成部であるのでマクロプランニング部は「机の前」Spaceオブジェクトを作り、「歩く」という命令をアニメーション生成部に送る。この場合もマクロプランニング部はエージェントが移動する先の具体的な座標を計算する必要はなくSpaceオブジェクトの組を設定するだけでよい。アニメーション生成部ではSpaceオブジェクトから移動する目的地として最適な一点を受取り、アニメーションを出力する。

4.1 Space オブジェクトの具体例

本稿ではまず仮想世界のシーンを部屋の内部に限定し、その中で移動する、物を運ぶ、といった動作を行うことを考えた。そのために今回は次のSpaceを実装した。

- (1) 前後左右
- (2) 物体の上、物体が占める空間
- (3) 物体のそば、近く
- (4) 物体にエージェントの手が届く範囲

4.2 方向についての表現

ある物体を基準とし、方向を表す言葉で空間を指定する場合について考える。本稿ではその中で「前後左右」を取り扱う。ポテンシャルを考えるための基準として、物体は前後左右の方向軸を持つとする。方向軸の決定は観察者の位置や物体の性質に依存し、自己参照軸、対面鏡映軸、物体固有軸、物体固有軸に場合分けが可能である。どの軸を用いるべきかという判定はマクロプランナが判断する。⁹⁾

4.2.1 自己参照軸

自分自身を基準とする方向軸を自己参照軸という。前後左右の定義は図2(a)のようになる。人間が自分自身

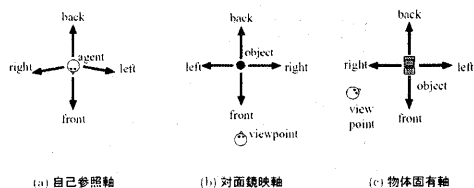


図2 方向軸の種類

を基準として方向を表す際には左右を表す方向軸は前後

を表す方向軸に垂直ではなく、10度ほど前にずれていることが実験により確かめられている¹²⁾ ことに基づき、本稿でも左右の方向軸をずらしている。前後左右を表すポテンシャル関数はそれぞれ式 (1) である。

$$f = \exp\left(\frac{-d_1^2}{(bd_2)^2}\right) \times \frac{1}{1 + e^{-ad_2}} \quad (1)$$

d_1 は注目している方向軸からの距離、 d_2 は観察者の中心を 0 とした方向軸方向の値、 a, b は定数パラメータである。

4.2.2 対面鏡映軸

観測者に対面した物体を中心に視点から見た方向軸を対面鏡映軸という。定義は図 2(b) で示される。

人が物体を用いて方向を決定する際、物体の性質を無視して決定しているとは考えにくい。例えば長方形のテーブルを見たとき、無意識に長方形の形に沿った方向軸が用いられる。その様子を図 3 に示す。Ball は机の影響を受けていない方向軸で判断するときには、左を示す方向ベクトルにもっとも近いので「table の左にある」と表現されるが、実際には机の影響を受けて「table の前にある」という表現のほうが自然である。本研究で

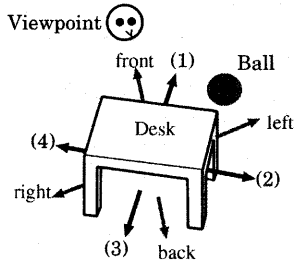


図 3 物体の外形に影響を受ける例

はすべての物体を長方形または円形に近似することにより、対面鏡映軸を調整する。物体が長方形と円形のどちらに近似されるべきかは頂点データから得られる第一、第二主成分ベクトルの長さの比を用いて決定する。対面鏡映軸の場合参照物体の用途、機能は考慮には入れない。軸の定義は図 4 のようになる。第一主成分と第二主成分はともに参照物体の中心を通り、前後左右を決める軸となる。「前」に対応するポテンシャル関数は式 (2) の

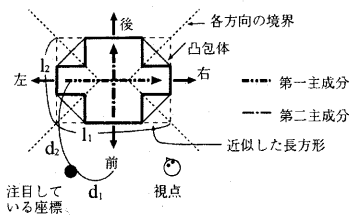


図 4 固有ベクトルを用いた対面鏡映軸

ようになる。

$$f = \exp\left(\frac{-d_1^2}{(d_2 - \frac{l_2}{2} + \frac{l_1}{2})^2}\right) \times \frac{1}{1 + e^{-ad_2}} \quad (2)$$

d_1 は前方向軸からの距離、 d_2 は物体の中心を 0 とした方向軸方向の値、 l_1, l_2 は長方形の辺の長さ、 a は定数パラメータである。積の前半部のガウシアン関数は方向を表す軸からの距離をパラメータとして、軸から離れるほどポテンシャルが低くなることを表している。 l_1, l_2 が示す辺は「前」「後」「左」「右」によって異なる。外形が円形に近似できるときは $d_1 = d_2 = 0$ とする。近似した長方形の頂点を通り、辺と 45 度で交わる直線上で第一項の値は e^{-1} となり、ここは「前」「後」「左」「右」のポテンシャルの大きさが入れかわる境界となる。ガウシアン関数のみでは、指示された方向と逆の方向について区別がつかないため、シグモイド関数との積を取ることで参照物体の中心よりも逆の空間のポテンシャルを 0 に減衰させる。

「机の前」の対面鏡映軸によるポテンシャル場を図 5 に示す。

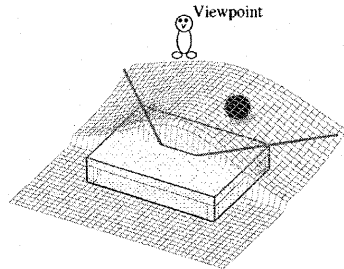


図 5 「机の前」を表すポテンシャル場

4.2.3 物体固有軸

オブジェクトには、その機能などにより前後左右の軸を独自に持っているものがある。そのような方向軸を物体固有軸という (図 2(c))。これは視点の位置に左右されない。一般にテレビは映像が映る面が前であり、車は進行方向が前である。このような軸は人の持つ経験により決定され、幾何的情報のみから決定するのは困難であるため、便宜的にすでに知っているものとする。このため、対面鏡映軸と異なり視点の位置にかかわらず軸は一意に決定される。方向軸が固定される点以外は対面鏡映軸と同じであり、ポテンシャル関数は式 (2) である。

4.3 位置についての表現

物体の位置を制約する空間表現を位置指示語と呼ぶ。本稿ではそのうち「上に乗っている」という位置指示語とエージェントが物体の位置を知るための「物体が存在する空間」を定式化した。ここでは「物体が存在する空間」を表す Space について説明する。

物体が存在する空間は物体の輪郭の内部であると考えられる。輪郭はオブジェクトの凸包体で近似する。ポテンシャルは式 (3) で定義する。

$$f = \begin{cases} e^{-\frac{d_1^2}{l^2}} & : \text{輪郭の内側} \\ e^{-\frac{d_1^2}{l^2}} \times e^{-\frac{d_2^2}{a}} & : \text{輪郭の外側} \end{cases} \quad (3)$$

d_1 は参照物体の重心からの距離、 d_2 は参照物体の最も近い凸包体の辺からの距離、 l は参照物体の重心から最も近い頂点までの距離、 a は定数パラメータである。参照物体の重心が最も存在の可能性が高く、重心から離れるに従ってポテンシャルは減少する。凸包体の外側では物体は存在しないのでガウシアン関数との積をとるよりポテンシャルを 0 に減衰させる。

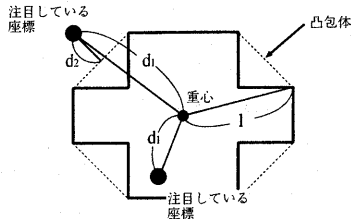


図6 物体の存在する空間

4.4 距離についての表現

物体の距離関係に注目した空間指示語を距離指示語と呼ぶ。今回は「そば」「近く」の定式化を行う。「そば」と「近く」の違いは参照物体からの距離の違いである。「そば」は「近く」よりも参照物体に近い位置を表現しており、ポテンシャル関数は式(4)のように定義する。

$$f = e^{-\frac{(d-l)^2}{ah}} \quad (4)$$

d は参照物体の最も近い輪郭からの距離、 h は参照物体の高さ、 l は指示された離れるべき距離、 a は定数パラメータである。参照物体の輪郭からの距離に応じてポテンシャルが変化する。「そば」では $l=0$ とすることで輪郭に最も近い位置でポテンシャルが最大となるようにする。「近く」では $l=h$ とし、輪郭から参照物体の高さと同じ距離だけ離れた地点でポテンシャルが最大となる。また、任意の距離での指定にも対応が可能である。その様子を図7に示す。

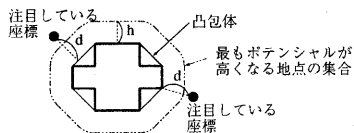


図7 参照物体から一定距離離れた地点

4.5 行動するための制約

エージェントが物を持つためには、手が届く範囲にすることが必要となる。そのため、本稿では「物体に手が届く」空間を Space オブジェクトとして定式化した。

物体を持つ際、どの部分を持つか、どのように持つかなど、解決しなければならない問題が多数あるが本稿では簡単化のため空間内で持つことのできる物体を球に限

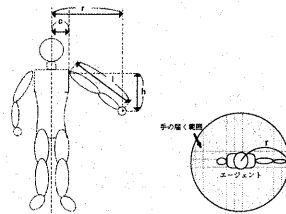
定し、エージェントはそれに触れることで持つことができるとした。

ポテンシャル関数は式(5)のように定義した。

$$f = e^{-\frac{(d-\frac{c}{2})^2}{a}} \times \frac{1}{1 - e^{b(d-r)}} \quad (5)$$

$$\text{但し } r = c + \sqrt{l^2 - h^2} \quad (6)$$

d はオブジェクトの中心から注目している座標までの距離、 c はエージェントの体の中心から肩までの距離、 l はエージェントの手の長さ、 h は肩の高さとオブジェクトの中心の高さの差、 a, b は定数パラメータである (図8)。



(a) 真横から見た場合

(b) 真上から見た場合

図8 エージェントの手が届く範囲

5. Space を利用したシステムとその動作結果

「青い机の前に行くって」という命令を実行するための必要な目的地の導出過程を図9、10、11、12、13に示す。「青い机の前」に移動するために必要な Space は3つの Space を合成することで構成される。まず、対面鏡映軸による「青い机の前」Space を考える (図10)。「前」という制約のみでは距離に関する制約が無く、机から離れた位置でもポテンシャルが高くなっている。机に近いほど空間ほど机との関係性が強調されると考え、距離に関する制約が存在しない場合は「青い机のそば」Space を合成し、机に近づくほどポテンシャルを大きくする。距離に関しての制約が存在するときには別の Space を合成する。(図11)。最後にエージェントが机にめり込むことを防ぐため、「青い机が存在しない」Space を合成する (図12)。以上の操作によりできる Space のポテンシャル関数から命令を満たすための一点を導出することでエージェントは移動の目的地を計算することができる。

6. おわりに

本稿で作成した Space を用いることにより、自然言語による指示文中の空間的制約の記号的取扱いと具体的な座標の導出が容易に行えるようになった。これによりマクロプランニング部では仮想世界の状況を記号表現として扱ったまま具体的な座標を考慮することなくプランを評価でき、アニメーション生成部では自然言語に関わり無く Space の持つポテンシャル場から座標を計算し、そ

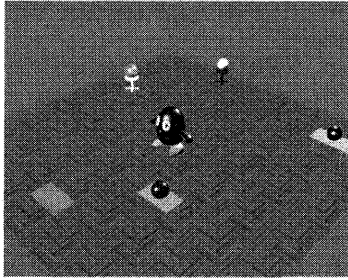


図 9 シーンの初期状態

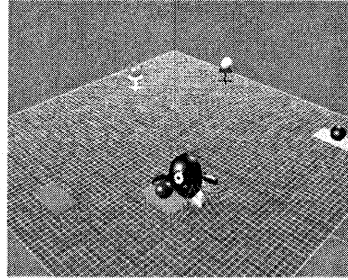


図 13 移動結果

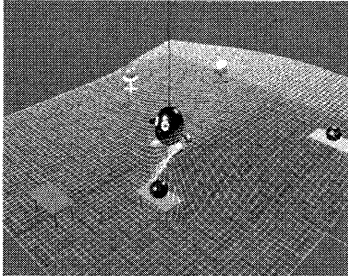


図 10 「青い机の前」のポテンシャル場

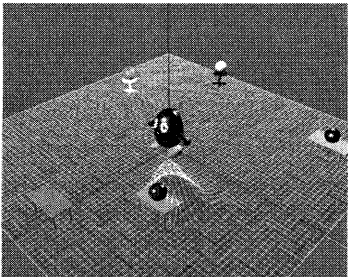


図 11 「青い机の前」×「机のそば」のポテンシャル場

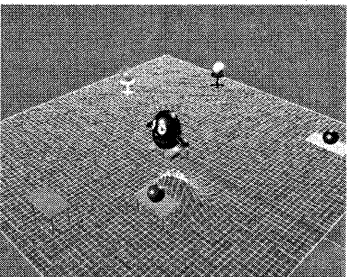


図 12 「青い机の前」×「机のそば」×「青い机が存在しない」のポテンシャル場

れを用いることができる。

Space オブジェクトではポテンシャル関数を用いることで空間表現の概念や、空間的制約を統一的に扱うことを可能にした。今後は Space オブジェクトを利用することで、今回のシステムでは厳密に取り扱っていない物体

の掴む際の位置を導出するなど、Space の種類を増やすことで動作の完成度を高め、様々な指示に対応できるようにする必要がある。

参 考 文 献

- 1) J.Cassell.: "More than Just Another Pretty Face: Embodied Conversational Interface Agents", Communication of the ACM, Vol.43, No.4, pp.70-78, 2000 年
- 2) N.I.Badler, M.S.Palmer, and R.Bindiganavale.: "Animation control for realtime visual humans.", Communication of the ACM, Vol.42, No.8, pp.64-73, 1999 年
- 3) R.Bindiganavale, W.Schuler, J.Allbeck, N.Badler, A.Joshi, and M.Palmer. "Dynamically Altering Agent Behaviors Using Natural Language Instructions.", Autonomous Agents 2000, pp.293-300, June 2000
- 4) 山田、西田、堂下: "二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 年
- 5) 山田、網谷、星野、西田、堂下: "自然言語における空間描写の解析と情景の再構築", 情報処理, Vol.31, No.5, pp.660-672 1990 年
- 6) 山田 篤、西田 豊明: "自然言語処理における図的表現の利用", 人工知能学会誌, Vol.9, No.2, pp.201-204 1994 年
- 7) 新山 雄介、秋山 英久、鈴木 泰山、徳永 健伸、田中 穂積: "自然言語を理解するアニメータッドエージェントのための 3 次元仮想空間における位置の表現と処理", 第 13 回人工知能学会全国大会論文集, pp.217-220 1999 年
- 8) 八嶋 絵美子、齋藤 豪、奥村 学、中嶋 正之: "仮想空間における位置関係の表現に関する研究", 電子情報通信学会 総合大会, 2002 年 3 月, D-5-11
- 9) 八嶋 絵美子、齋藤 豪、奥村 学、中嶋 正之: "仮想空間における位置関係の表現", 情報処理, Vol.2001, No.89, pp.9-12
- 10) 田中 茂範、松本 曜: "空間と移動の表現", 研究者出版, 1997 年
- 11) A.Herskovits.: "Language and Spatial Cognition.", Cambridge University Press, 1986 年
- 12) 中西 卓哉、今井 むつみ、石崎 俊: "仮想空間における空間指示の研究", 電子情報通信学会 研究会資料, NLC-97-3, 1997 年