

Double Cross Model に基づく相対位置情報アノテーション

川端 良子^a・大村 舞^a・浅原 正幸^a・竹内 誉羽^b

Spatial ML や ISO-Space など、言語が表現する位置情報を記述する方法が提案されている。これらは固有位置情報や絶対位置情報（東西南北）を記述するのに有効であるが、対話の中で多用される一人称視点の相対位置情報（前後左右）を記述するには適していない。相対的な参照表現の曖昧性解消するには、単に1つの有向辺のみによる表現は本質的に不十分で、実体の向きを含んだフレームとしての2つ以上の有向辺を用いる必要がある。一方、空間論理の分野では、Double Cross Model は空間論理の分野で3点の相対的な位置情報を表現するために提案された。本研究では、Double Cross Model を用いて対話の中の相対的な参照表現を形式化し、アノテーションを行ったので報告する。

キーワード：位置情報, 空間論理, 参照表現

Spatial Information Annotation Based on the Double Cross Model

YOSHIKO KAWABATA^a, MAI OMURA^a, MASAYUKI ASAHARA^a and JOHANE TAKEUCHI^b

Spatial ML and ISO-Space have been proposed as methods to describe spatial information that appears in a language. Although these methods are effective for describing proper position information and absolute references (such as cardinal directions), they are not suitable for describing the relative references (such as front, back, left, and right) inherent in dialogue as entities in space. Ambiguity in relative references cannot be resolved using only directed edges (ordered pairs), and it is necessary to maintain two or more directed edges as frames, including the orientation of the entities. The double cross model is used in spatial logic to resolve the ambiguity of location information. In this study, we attempted to represent relative references in dialogue using the double cross model and present our findings.

Key Words: *Spatial Information, Spatial Logic, Reference Expressions*

^a 国立国語研究所, NINJAL, Japan

^b ホンダ・リサーチ・インスティテュート・ジャパン, HRI-JP

本論文は、「Double Cross Model による位置情報フレームアノテーション」言語処理学会第 29 回年次大会 (NLP2023) 川端・大村・浅原・竹内 (2023), “Spatial Information Annotation Based on the Double Cross Model”, Pacific Asia Conference on Language, Information and Computation (PACLIC 37), Kawabata, Omura, Asahara and Takeuchi (2023) を元に修正したものである。本研究では、企業内で作成されたデータについて、部分的に開示できる情報を、応用システム論文として共有するものである。

1 はじめに

他の人にどこかへ連れて行ってもらいたい時, 住所・緯度・経度情報などの固有位置情報や東西南北などの絶対位置情報のみならず, 前後左右などの話者の一人称視点からの相対位置情報によって位置情報を共有することが多い. このような相対位置情報を表現する場合, 単一の有向辺のみに基づく抽象化では, 本質的に位置情報の曖昧性解消ができない場合がある.

例えば, 「東京タワーを背にして右前方に豆腐屋があります」という表現を有向辺で表現するためには2つの辺を利用する必要がある. 1つはランドマークである「東京タワー」と空間的実体である「話し手」の現在の位置との相対的な配置を表し, もう1つは「話し手」の現在の位置とランドマークである「豆腐屋」の位置との相対的な配置を表す. 正確な位置情報を記述するためには, これらの2つの有向辺の情報を単一の位置情報フレームとして保持し, 3点の相対位置を定義することが必要である.

空間論理の分野において, 相対位置情報を表現する Double Cross Model が提案されている. Double Cross Model は文字通り2つの十字(cross)を使用する位置情報フレームである. 2つの十字の中心に2つのランドマークもしくは空間内実体を配置したうえで, 3つ目のランドマークもしくは空間内実体の相対位置情報を表現することができる.

図1に例を示す. 表現「[東京タワー](id:T1)を背にして右前方に[豆腐屋](id:T3)があります

【作例】
 [東京タワー](id:T1)を背にして右前方に[豆腐屋](id:T3)があります

※ (id:H0) は聞き手を表す

図

```

            [ id  T1
              type landmark
              slot 15
            ],
            [ id  T3
              type landmark
              slot 12
            ],
            [ id  H0
              type se
              slot 14
              dir  1
            ]
        
```

AVM

```

            [
              { "id" : "T1",
                "type" : "landmark",
                "slot" : 15},
              { "id" : "T3",
                "type" : "landmark",
                "slot" : 12},
              { "id" : "H0",
                "type" : "se",
                "slot" : 14,
                "dir" : 1}
            ]
        
```

JSON

```

            - id: T1
              type: landmark
              slot: 15
            - id: T3
              type: landmark
              slot: 12
            - id: H0
              type: se
              slot: 14
              dir: 1
        
```

YAML

注: 本フレームに必要なラベルと識別子 (T1, T3) のみを示す。

図 1 Double Cross Model に基づく相対位置情報表現の例

す」に対して、聞き手の id を H0 としたうえで、図 1 の左下の二つの十字の図のように表現する。

本論文では、位置情報を共有する対話中の相対位置情報を、Double Cross Model を利用して表現することを提案する。さらに、本研究では、絶対位置情報・方向・向き・時間的距離・空間的距離・位相情報（部分・全体情報）のフレーム情報の付与手法について示す。

本論文の構成は以下の通りである。2 節では、空間情報フレームのアノテーション手法について解説する。3 節では、提案手法である Double Cross Model に基づく相対位置情報アノテーション手法を示す。4 節では、実際のアノテーションを行ったデータの統計について示す。5 節に、まとめと今後の展開について示す。

なお本研究は、ホンダ・リサーチ・インスティテュートにおいて、対話データの収録・転記・さまざまなメンションのタグ付けされたものに対して、どのように相対位置情報を適切にタグ付けするかについて提案するものである。

2 空間情報フレームのアノテーション手法

本節では、空間論理の考え方にに基づき、空間情報フレームのアノテーション手法を提案する。Renz らの空間論理の解説 (Renz and Nebel 2007) には、区間代数からはじめ、2.2 節に示す位相に基づく抽象化を解説するとともに、向きを入れる手法としての Double Cross Model を示している。以下では、Renz らの定義に基づき、空間情報の抽象化について解説する。さらに、言語表現に対する空間情報アノテーション手法として、Spatial ML (Mani et al. 2008) や ISO-space (Pustejovsky and Yocum 2014; Pustejovsky 2017) などの空間情報フレームアノテーションの主要な概念について解説する。

2.1 空間情報の抽象化

本節では空間情報の抽象化手法について解説する。本研究で扱う表現は以下のとおりである：

- ランドマーク (Landmarks)：緯度や経度、または住所など、特定の位置情報が定義される場所。
- 空間内実体 (Spatial Entities)：空間に位置する実体。話し手・聞き手もこれに含まれる。
- シグナル表現 (Signals)：位置情報、方向情報、距離情報などを示す表現。

位置情報の形式化については、その参照表現の型として 3 種類あるとされる：

- 固有 (Intrinsic)：場所に内在する固有の向き・位置。
- 絶対 (Absolute)：鳥瞰による指示参照（東西南北）。
- 相対 (Relative)：実体からの視点に基づく指示参照（前後左右）。前提として、視点を持つ主体が向き (Orientation) を持つ。

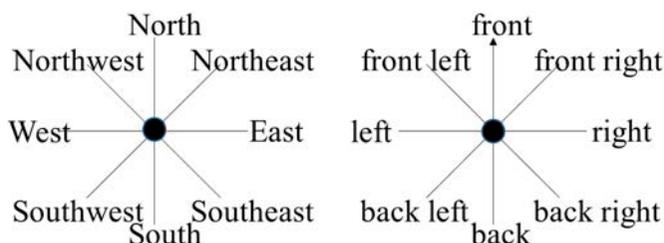


図 2 Project-based model (左: 絶対 (absolute), 右: 相対 (relative))

空間内の位置に基づき方向を規定する方法として project-based model (図 2) (Ligozat 1998) がある。絶対位置情報表現はある出発点からの目標点の位置・方向を鳥瞰視点に基づき「東西南南北」を用いて示す。相対位置情報表現はある出発点からの目標点の位置・方向を空間内の一人称視点に基づき「前後左右」を用いて示す。project-based model では位置・方向を有向辺にて表現する。

言語表現が絶対的な指示参照しか含まない場合、その指示参照は固有位置情報を参照点とした上記の有向グラフモデルを使用して位置情報を抽象化することができる。しかし、実際には、空間内実体である話し手と聞き手は空間内に存在し、それぞれの視点からの相対位置情報表現を使用して位置情報を伝達する傾向にある。特に必ずしも空間内の絶対位置表現 (東西南南北) が対話参加者で共有されるわけではなく、相対位置情報表現を使わざるを得ない状況がある。

しかしながら、既存の project-based model による相対位置情報表現の形式化は空間内実体がどちらを向いているのかという情報を適切に保持しない。3 節に示す Double Cross Model を用いた相対位置情報アノテーション手法では、3 点の相対位置情報を保持することで適切に相対位置情報表現を抽象化する。

2.2 位相情報の抽象化

位相情報とは空間内物体の包接関係を抽象化したものである。ランドマークや空間内実体には大きさがあり、その包接関係は位置情報のなかでも重要な情報である。Spatial ML や ISO-Space などの先行研究における言語表現の空間情報の抽象化 (Mani et al. 2008; Pustejovsky and Yocum 2014; Pustejovsky 2017) は主に位相情報に基づいた抽象化である。これは、移動表現などの述語がランドマークと空間内実体の相対的な位相変化の意味を表出するため、述語の事象意味論として位置情報の抽象化を目指したためである。

位相情報として、部分全体関係や境界情報を定義する手法として RCC-8 (Region Connection Calculus-8) (Randell et al. 1992) が提案されている。文献によっては RCC8 とも記されるが、本稿では本文中で言及する場合 RCC-8 に統一し、ラベル名として rcc8 を用いる。以下の RCC-8

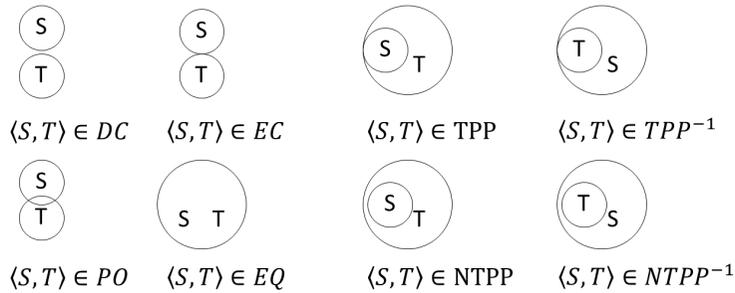


図 3 位相情報 (RCC-8)

で規定する 8 種類の関係 (図 3) を示す.

- DC: Disconnected (非連結)
- EC: External connection (外接)
- PO: Partial overlap (部分一致)
- EQ: Equal (完全一致)
- TPP: Tangential proper part (被内接: source が target に接した部分集合)
- TPP^{-1} : Inverse of TPP (内接: target が source に接した部分集合)
- NTPP: Non-tangential proper part (被内包: source が target に接しない部分集合)
- $NTPP^{-1}$: Inverse of NTPP (内包: target が source に接しない部分集合)

Pustejovsky (2017) の ISO-Space における Motion タグのクラスは, 移動表現の位相の変化を分類したものである. 表 1 に Motion タグの移動表現の位相の変化を RCC-8 で表現したものを示す. 表中 IN は TPP と NTPP の集合和を表す. 例えば, Leave は, 何らかの空間内実体の内側 ([IN|EQ]) から, 部分一致・外接 (PO, EC) を経て, 空間内実体の外部 (DC) への移動を表す. Reach は, Leave の逆順の移動を表す. Hit は, 空間内実体の外部 (DC) から外接 (EC), さらに場合によっては部分一致 (PO) の移動を表す. このように動詞が表出する位相の変化の記述を目的として, 事象意味論の研究が進められている.

我々の提案手法でも非連結 (DC) 以外の位相情報は同様のラベルを用いて表現を行う.

3 Double Cross Model に基づく相対位置情報アノテーション

以下では, 本研究で提案する相対位置情報アノテーションを示す. 通常の位置関係は 2 地点のものをリンク (有向辺 or 無向辺) で表現する. 本研究では 3 地点以上の情報をフレーム (Minsky 1974) として表現する. 本稿では, 3 地点以上の位置関係を表現するために Double Cross Model 上に配置して抽象化するとともに, 位相や距離の情報を表す抽象化としてのフレームを提案す

Motion Class	Examples	RCC-8
Move	run, fly, drive	未定義
Move_External	traverse, pass	(DC, DC, DC) or (EC, EC, EC)
Move_Internal	walk around	([IN EQ], [IN EQ], [IN EQ])
Leave	leave, depart	([IN EQ], (PO, EC), DC)
Reach	arrive, enter, reach	(DC, (EC, PO), [IN EQ])
Detach	take off, pull away	(PO, EC, DC) or (EC, DC)
Hit	land, hit	(DC, EC, PO) or (DC, EC)
Cross	cross	(DC, (EC, (PO, [IN EQ], PO), EC), DC) or (EC, (PO, [IN EQ], PO), EC) or (PO, [IN EQ], PO) or (TPP, NTPP, TPP)
Follow	follow, chase	(IN, NTPP, IN)
Deviate	flee, run from	(IN, EC, DC)
Stay	remain, stay	変化なし

注：() 内の左から右の順に位相が変化する。[] はいずれかの位相を表す。

表 1 ISO-Space Motion タグにおける位相情報 (Pustejovsky (2017) より)

る。本研究において、フレームは3地点以上の位置関係を矛盾なく表現できるものを単位とする。矛盾なく説明可能な場合には、4地点以上の位置関係を一つのフレームとして表現することを許す。

まず、アノテーションの形式について、図1を用いて示す。次に先行研究にならい位相情報の扱いについて示す。さらに Double Cross Model を用いた相対位置情報アノテーション手法を示し、同型の扱いについて解説する。最後に距離情報の扱いについて示す。

3.1 アノテーションの形式化

アノテーションの形式について、図1の例を用いて示す。アノテーションは、位置情報に関するフレーム知識をハッシュの順序なしリスト(集合)として記述する。

まず、テキスト中の連続文字列を抜き出して、その文字列を指し示す識別子(id)を付与する。そのidに対して、型(type)情報を付与する。ランドマークの場合にはtype=landmarkを、空間内実体の場合にはtype=seを付与する。図1の例では、メンション「東京タワー」にid=T1を付与し、「豆腐屋」にid=T3を付与する。このメンションの切り出し作業は本研究の上流工程で行われ、提供されたものをそのまま利用した。

また、本研究では新たに空間内実体の話し手を導入する。空間内実体の聞き手としてid=H0を設定する。

以下に記述する内容について示す：

- id: (mention id)

テキスト内のメンションの識別子。これは、位置表現もしくは手がかり句に対して割り

当てられる。さらにテキスト中に出現しない空間内実体として、対話の参与者である S0 (話し手) および H0 (聞き手) を特別の識別子として用いる。

- **type**: [landmark, se, signal, distance, topology]
 テキスト内のメンションの型。もしくはフレームの型。次のいずれか：
 - **landmark**: 位置情報 (動かないもの)。
 - **se**:
 空間内実体 (動くもの)。話し手 (S0) や聞き手 (H0) を含む。
 - **signal**: 相対位置を示す手がかり句。
 - **distance**: 距離を表すフレーム
 - **topology**: 位相表現
- **rcc8**: [EC, PO, EQ, TPP, NTPP]
 位相情報を表すラベルを付与する。外接 (EC)・部分一致 (PO)・完全一致 (EQ)・被内接 (TPP)・被内包 (NTPP) のいずれか。
- **src**: (mention id)
 位相情報・距離を表す起点となるメンションの ID。外接 (EC)・部分一致 (PO)・完全一致 (EQ) は対称関係なので、次の終点となるメンションの ID **tgt** と入れ換えても成立するが、いずれかの順序のもののみを記述する。被内接 (TPP)・被内包 (NTPP) は非対称関係である。内接 (TPP^{-1})・内包 ($NTPP^{-1}$) については、それぞれの対称関係にある被内接 (TPP)・被内包 (NTPP) の **src** と **tgt** を入れ換えて記述する。
- **tgt**: (mention id)
 位相情報・距離を表す着点となるメンションの ID。
- **slot**: [1-15]
 メンション ID の位置情報・空間内実体の相対位置を示す。以下に示す Double Cross Model におけるスロット番号。
- **dir**: [1-15]
 メンション ID の空間内実体の向いている方向を示す。以下に示す Double Cross Model におけるスロット番号。
- **absdist**: (距離情報・時間情報)
 絶対距離情報。到達に要する時間による距離表現を許す。
- **reldist**: (距離情報)
 相対距離情報。14 と 15 の間の距離を 1 とした場合の距離。

これらの付加情報は、属性値行列 (AVM: Attribute Value Matrix) 形式で保持する。さらに、ID 間の関係 (2 項関係) やフレームについても属性値行列で記述する。結果として、アノテーションは属性値行列の集合をなす。属性値行列は、計算機上では JSON もしくは YAML 形式

で保持する.

なお, 絶対位置情報による参照表現の場合には, 必ずしも Double Cross Model を用いる必要はない. しかしながらアノテーションの一貫性のために, 本研究では Double Cross Model を用いて, 絶対位置情報も表現する.

3.2 位相情報アノテーション

本研究においても RCC-8 (図3) に基づき位相情報をアノテーションする. 以下では, 実際のアノテーション事例により, アノテーション手法について説明する.

図4に被内包 (NTPP) のタグ付け例を示す. 例文中に `type=landmark` のメンションとして「公園」(`id=T412`)と「マンション街」(`id=T413`)が定義されている. この位相情報を RCC-8 の起点を `src=T412`, RCC-8 の着点を `tgt=T413` としたうえで, RCC-8 の関係を `rcc8=NTPP` と付与する.

図5に完全一致 (EQ) の例を示す. 「アニバーサリー」(`id=T110`)と「お店」(`id=T111`)は同一実体であり, この位相情報として, RCC-8 の起点を `src=T110`, RCC-8 の着点を `tgt=T111` としたうえで, RCC-8 の関係を `rcc8=EQ` と付与する.

なお, 図6のように複数の同一指示参照表現が出現する場合, 同一指示参照表現のすべての組み合わせについて付与するのは煩雑である. このような場合には, 同一指示参照表現のうちで最も近いメンションの対についてのみ関係を付与する. 図6の例では, 最も近いメンション対である (T366, T367), (T367, T369), (T369, T370) には関係を付与するが, そうではない (T366, T369),

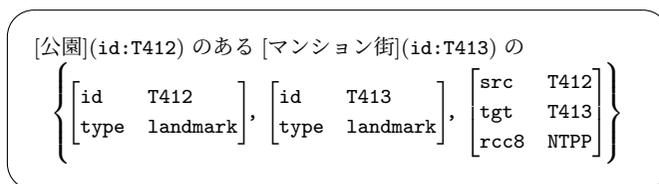


図4 位相情報アノテーション例 (NTPP)

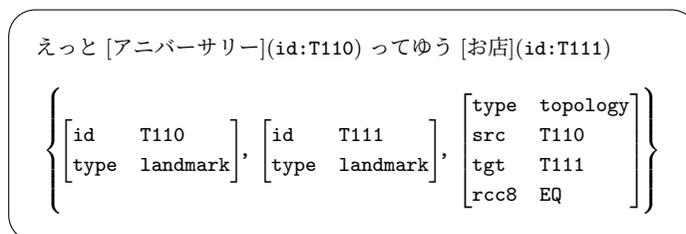


図5 位相情報アノテーション例 (EQ)

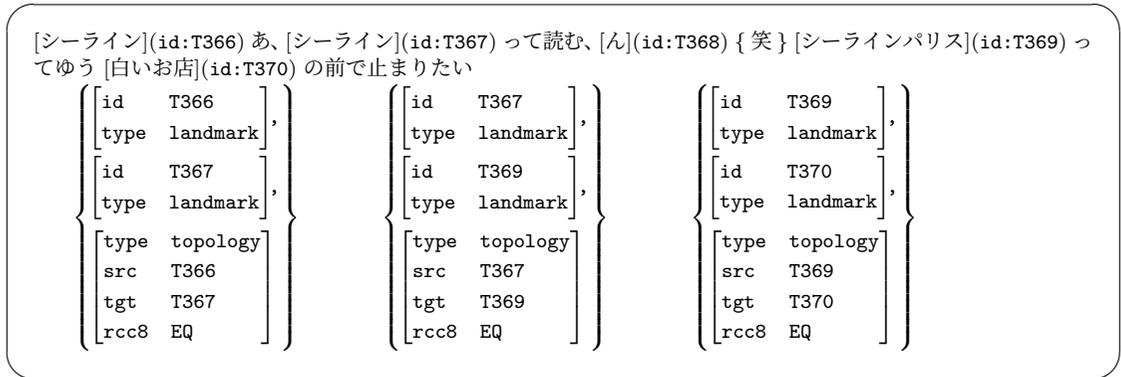


図 6 位相情報アノテーション例 (EQ の連続)

(T366, T370), (T367, T370) には関係を付与しない。最も近いメンション対であるか否かは同一指示参照表現の中のみで判定する。図 6 において切り出されているフィルター表現 T368 は {T366, T367, T369, T370} と同じ実体を参照しないために、最も近いメンションの候補として考慮しない。

本研究では、デフォルトを非連結 (DC) と仮定し、非連結の場合にはラベルを付与しない。また内接 $TPP^{-1} \langle S, T \rangle$ と内包 $NTPP^{-1} \langle S, T \rangle$ はラベルとして定義せず、被内接 $TPP \langle T, S \rangle$ や被内包 $NTPP \langle T, S \rangle$ のように起点と着点を反転することで表現し、ラベルの数を減らす。

RCC-8 は隣接しているか離れているかを表現することに重きを置いており、位置情報の特定のために必要な相対的な方向表現を適切に表現することができない。位相情報と相対位置情報を適切にアノテーションとして表現することが必要である。

3.3 Double Cross Model による相対位置情報アノテーション

Double Cross Model (図 7) (Freksa 1992) は、2つの空間内実体 (もしくはランドマーク) を図中の 14 (relatum) と 15 (origin) の位置に配置し、その向きを決めるための 3 つ目の要素 (referent) を図中の他の位置に配置することで、3 点で向きと相対位置を指定する表現手法である。

図 1 の例では、「東京タワー」(id=T1) を 15 に配置し、聞き手 (id=H0) を 14 に配置し、「豆腐屋」(id=T3) を 12 に配置することで、相対位置を示す。さらに、聞き手がどちらを向いているのかの情報を dir=1 として Double Cross Model を用いて表現する。

Double Cross Model においても、同一指示 (EQ) のメンションが複数出現する場合には、最も近いメンションについてのみ記述した。また Double Cross Model に基づく空間論理演算によっ

て類推される関係までは付与せずに, 言語が表現する配置をそのまま記述するのにとどめた.¹

3.4 Double Cross Model における同型

Double Cross Model は origin, relatum, referent の 3 項関係を表す. 本来の Double Cross Model においては, origin を 15 に配置し, relatum を 14 に配置したうえで, referent を他の位置に配置することで相対位置情報を表現する. Scivos and Nebel (2001) は, 図 8 の右のように {left, straight, right} × {forward, perpendicular, centre, line, back} と参照点の名前を定義した.

3 項の Double Cross Model では, 表 2 のように, 1 つの構造に対して異なる 6 つ (3-permutation) の表現が可能である. Zimmermann and Freksa (1996) は, この同型の表現について ‘inversion’

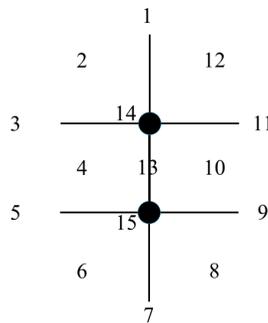


図 7 Double Cross Model

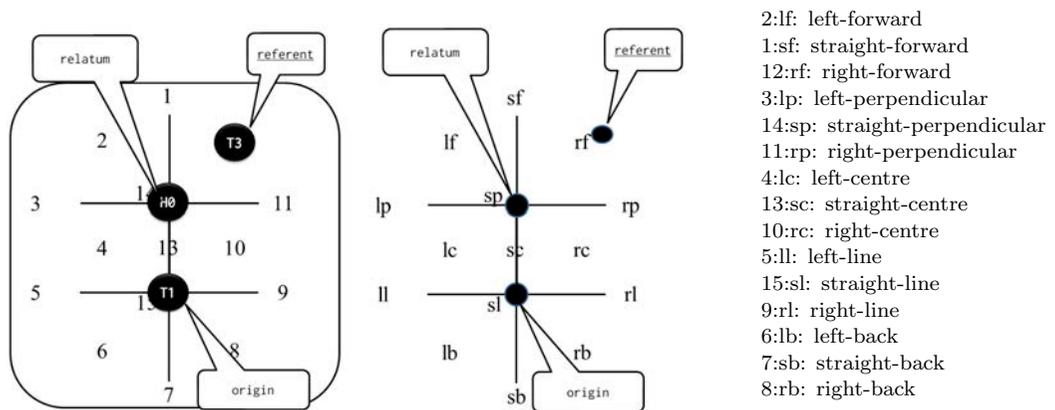


図 8 Double Cross Model における 3 項関係 : 〈 origin, relatum, referent 〉

¹ 例えば, a, b, c, d の 4 点があるとする. (a, b: c) について (15, 14: 1) の配置であり, (b, c: d) について (15, 14: 12) の配置の場合, (a, c: d) は空間論理演算により (15, 14: 12) の配置になることが推論されるが, そこまでの記述は行わない.

操作	origin, relatum; referent
identical	a,b;c
inversion	b,a;c
homing	b,c;a
inverse homing	c,a;b
short cut	a,c;b
inverse short cut	c,a;b

表 2 origin が a, relatum が b, referent が c の場合の順列に基づく同型

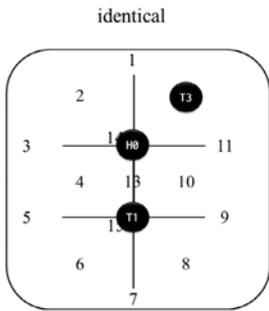


図 9 同型: inversion

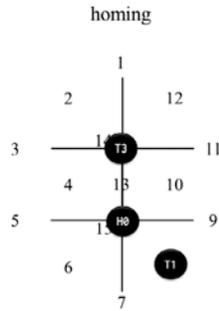
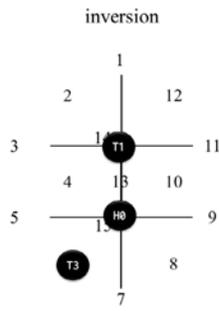


図 10 同型: homing, inverse homing

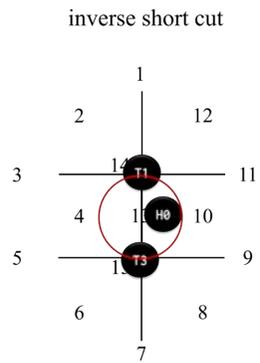
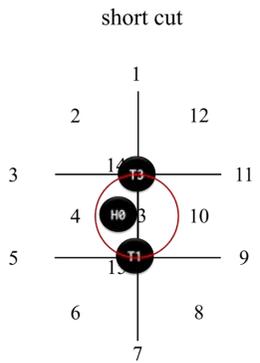
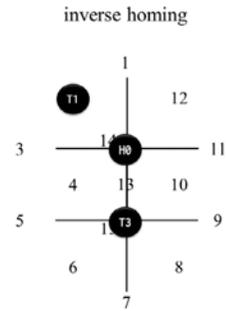


図 11 同型: short cut, inverse short cut

(図 9) , ‘homing’ (図 10) , ‘short cut’ (図 11) のように定義する.

‘inversion’ は origin と relatum を入れ換える操作である (図 9). origin を 14 に配置し, relatum を 15 に配置することで, 元の表現を 180 度回転させたものとなる.

‘homing’ は, 元の relatum を新しい origin に, 元の referent を新しい relatum に, 元の origin を新しい referent に入れ替える操作である (図 10). ‘inverse homing’ は ‘homing’ 後に ‘inverse’ したものである.

‘short cut’ は relatum と referent を入れ換える操作である (図 11). ‘short cut’ においては, 14 と 15 を直径とする円が重要な役割を果たす. 元の構造で origin が 15 にあり, relatum が 14 にある場合に, ‘short cut’ 前の referent が 2 か 12 にある場合には, 入れ換えた新 referent は必ず円の中にある. ‘short cut’ 前の referent が 4 か 10 にある場合には, 入れ換えた新 referent は必ず円の外にある. ‘short cut’ 前の referent が 3 か 11 にある場合には, 入れ換えた新 referent は必ず円の上にある (タレスの定理). ‘inverse short cut’ は ‘short cut’ 後に ‘inverse’ したものである.

タグ付け作業者は, この順列に基づく同型の異なる表現のいずれで記述してもよいこととする. 実際に機械処理する際に, この空間論理における操作 (inversion, homing, short cut) を適切に処理する.

3.5 距離情報フレーム

Double Cross Model は相対位置情報を記述するもので, 距離情報は記述しない. 言語表現には絶対距離情報表現 (Hernández et al. 1995; Clementini et al. 1997) と相対距離情報表現 (Islis and Moratz 1999) の 2 つに分けられる.

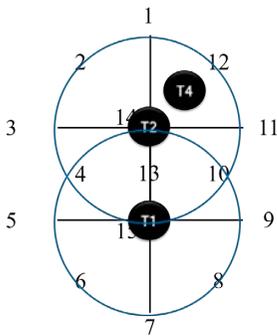
絶対距離情報表現は, 2 点間の距離を, m, km などの具体的な値や到達に要する時間で表現するものである. 図 12 の例では, 距離情報フレーム `type=distance` を定義し, 起点として `src=T13`, 着点として `tgt=T16`, そしてその 2 点間の絶対距離情報を `absdist=20m` として記述する. 今回タグ付けしたデータは 360 度カメラの静止画を刺激としたもので, 到達に要する時間で表現する距離表現は含まれなかった.

相対距離情報表現は, 2 点間の距離を, 別の 2 点間の距離の相対値により表現するものである. 今回タグ付けしたデータには, 相対距離情報表現は含まれていなかった. Double Cross Model における相対距離情報の表現手法として, ‘origin’ と ‘relatum’ の距離を基準として, そのいずれかの点と ‘referent’ との相対距離を記述する方法を考えていた. 具体的には, 図 13 のように, ‘origin’ と ‘relatum’ の 2 点によってなす辺を基準となる距離とし, 他の距離をその倍数で

[ジングウマエコウバン](id:T13) から [メイジジングウ](id:T14) のほうに [20メートル](id:T15)[行った歩道](id:T16) で [待ってて](id:T17)

$$\left\{ \begin{array}{l} \left[\begin{array}{ll} \text{id} & \text{T13} \\ \text{type} & \text{landmark} \\ \text{slot} & 15 \end{array} \right], \left[\begin{array}{ll} \text{id} & \text{T14} \\ \text{type} & \text{landmark} \\ \text{slot} & 1 \end{array} \right], \left[\begin{array}{ll} \text{id} & \text{T16} \\ \text{type} & \text{landmark} \\ \text{slot} & 14 \end{array} \right], \left[\begin{array}{ll} \text{id} & \text{T15} \\ \text{type} & \text{signal} \end{array} \right], \left[\begin{array}{ll} \text{type} & \text{distance} \\ \text{src} & \text{T13} \\ \text{tgt} & \text{T16} \\ \text{absdist} & 20\text{m} \end{array} \right] \end{array} \right\}$$

図 12 絶対距離情報表現



(作例) [タチカワエキ](id:T1) から歩いて [タカマツエキ](id:T2) まで進み、道なりに右前方に曲がって、そこから今まで歩いてきた [半分の距離](id:T3) だけ進むと [ララポートタチヒ](id:T4) に着きます

$$\left\{ \begin{array}{l} \left[\begin{array}{ll} \text{id} & \text{T1} \\ \text{type} & \text{landmark} \\ \text{slot} & 15 \end{array} \right], \left[\begin{array}{ll} \text{id} & \text{T2} \\ \text{type} & \text{landmark} \\ \text{slot} & 14 \end{array} \right], \left[\begin{array}{ll} \text{id} & \text{T4} \\ \text{type} & \text{landmark} \\ \text{slot} & 12 \end{array} \right], \\ \left[\begin{array}{ll} \text{id} & \text{T3} \\ \text{type} & \text{signal} \end{array} \right], \left[\begin{array}{ll} \text{type} & \text{distance} \\ \text{src} & \text{T1} \\ \text{tgt} & \text{T2} \\ \text{reldist} & 1 \end{array} \right], \left[\begin{array}{ll} \text{type} & \text{distance} \\ \text{src} & \text{T2} \\ \text{tgt} & \text{T4} \\ \text{reldist} & 0.5 \end{array} \right] \end{array} \right\}$$

図 13 相対距離情報表現

reldist として記述する方法である。この例では、T1 と T2 の距離を基準となる 1 とおき、T2 と T4 の距離を 0.5 と記述する。

4 アノテーションデータ

本節では、マイクロモビリティと位置情報を共有することを目的とした収録対話に対するアノテーションデータについて示す。本研究はホンダ・リサーチ・インスティテュート・ジャパンと国立国語研究所の共同研究プロジェクトにより実施した。なお本研究は、対話データの収録・転記・メンションアノテーションがなされたものに対して、どのように相対位置情報を表現するかについて提案するものである。

4.1 アノテーション対象データ

以下でホンダ・リサーチ・インスティテュートで構成されたアノテーション対象のデータについて開示可能な範囲の情報をまとめる。

収録対話は、1 件の 360 度カメラの静止画を見ながら、ロボット役（実験実施者）と位置情報を共有する設定で行った。静止画は、東京都の表参道および豊洲近辺で撮影したものを利用した。収録に際して、ロボット役（実験実施者）は制限されたテンプレートの発話のみが可能である。実験協力者側の発話のみをアノテーション対象とした。

2022 年の 8 月に 10 人の実験協力者から音声収録を行った。各実験協力者は練習として 1 セッション実施したうえで、データ収録のための 20 セッション（20 静止画）実施し、合計 200 対話を収録した。1 セッションは 150 秒を目標としたうえで 210 秒の時間制限を設け、目的とする位置情報の共有が達成するか、210 秒の時点でそのセッションを打ち切った。音声収録に際し、ホンダ・リサーチ・インスティテュート・ジャパンの研究倫理審査を受け認可された。音

声データを転記し, 場所・位置関係のメンションを付与した. データの概要については, 付録 A に示す.

4.2 位相情報・相対位置情報アノテーション

国立国語研究所において3節に示したアノテーション仕様を提案し, 国立国語研究所内でアノテーションを行った.

位相情報 (rcc8) のアノテーションについては, 有向辺に抽象化したうえで brat² を用いてタグ付けを行った. 図 14 にアノテーション画面を示す. 位相情報は共参照情報の一般化であるが, 被内接 TPP・被内包 NTPP のような非対照関係を含むために, 有向辺で記述した. 完全一致 EQ・部分一致 PO・外接 EC については, 有向辺で付与するが, 矢印はどちらを向いていてもよいこととした. また, 完全一致 EQ は同値関係であり, 実質的に同値類を規定するために, 同値類内のすべての組み合わせを付与することはせず, 同値類内の隣接する要素にのみに付与した. アノテーションデータを, AVM などの記述形式に変換して用いる.

Double Cross Model の記述は, brat 上ではできないために, brat に示される識別子情報を見ながら, タグ付け作業者が YAML 形式で記述した.

表 3 にアノテーションデータを示す. ‘Turns’ の列は, ロボットとの対話におけるターン数を示す. そのうち左の数値は実験協力者側の発話数 (UTurns) とし, 右の数値を全体のターン数とする. ‘EQ’・‘TPP’・‘NTPP’・‘EC’・‘PO’ および ‘rcc8 Total’ の列は RCC-8 に基づく位相情報の数を示す. ‘Double Cross Model’ の列は, Double Cross Model によって記述したフレーム数を示す. ‘目的達成’ の列は, 各協力者が 20 回のセッションのうち, どれだけの回で自分の位置情報を 210 秒以内に共有できたかを示す.

表中, 完全一致 ‘EQ’ はいわゆる共参照を示している. 位置情報を共有する際に同じランドマーク・空間内実体を繰り返し言い換える傾向があり, EQ が頻繁に出現する. 被内接 TPP およ

19	System : どこに向かいますか								
20	User :	Fil	Loc (r)	NTPP	Fil	Loc	ACT		
21	System :								
22	User :	Fil	Loc (r)	NTPP	Fil	Fil	Loc (r)	Act	Fil

図 14 brat による rcc8 のアノテーション

² <https://brat.nlplab.org/>

Subject	Turns UTurns/全体	EQ 完全一致	TPP 被内接	NTPP 被内包	EC 外接	PO 部分一致	rcc8 Total	メンション 頻度	rcc8 /メンション 平均	メンション /UTurn 平均	Double Cross Model	目的達成
S01	72/146	99	14	6	29	1	149	226	1.3	3.1	0	15/20
S02	77/158	40	24	6	41	0	111	183	1.2	2.4	5	20/20
S03	78/159	31	1	8	23	0	63	132	1.0	1.7	7	18/20
S04	88/178	17	8	5	25	0	55	130	0.8	1.5	2	17/20
S05	72/144	34	4	1	13	0	52	107	1.0	1.5	0	19/20
S06	114/232	111	13	28	18	0	170	275	1.2	2.4	3	17/20
S07	56/114	27	19	25	25	0	96	156	1.2	2.8	1	19/20
S08	64/129	50	8	3	8	0	69	116	1.2	1.8	0	19/20
S09	91/183	24	5	2	6	0	37	129	0.6	1.4	3	18/20
S10	62/125	43	15	2	20	0	80	135	1.2	2.2	0	19/20
Total	774/1568	476	111	86	208	1	882	1589	—	—	21	181/200

表 3 アノテーションデータの統計: RCC-8 ラベル と Double Cross Model フレームの度数

び被内包 NTPP は部分全体関係を示し、主として建物内の店舗を参照する際に使用される。これらは外部に面しているかどうかで区別され、TPP は外部に面していることを示し、NTPP は外部に面していないことを示す。外接 EC は主に隣接するランドマーク・空間内実体に用いられる。部分一致 PO は部分的に重なる実体を表現するために利用されるが、都市部ではあまり部分的に重なるランドマークや空間内実体が少ないために数が限定的である。

Double Cross Model を用いて表現すべき言語現象は、10 人中 6 人の実験協力者が利用し、わずか 21 発話だった。本実験のように 360 度カメラの静止画の情報のみでは、絶対位置参照情報である東西南北が利用できない。このために本タスクでは相対位置参照情報のみで適切に位置情報を表現する必要があり、ランドマークの特徴とそのランドマークの隣接関係を含む位相情報に頼る傾向が確認された。

図 15 に各セッションに要した時間（実験協力者 10 人 × 20 セッション）を示す。棒グラフの下に表示されている数字は、Double Cross Model が用いられた表現の数を示す。白は目的達成したセッションを、黒は目的達成しなかったセッションを表す。Double Cross Model の表現を含む 17 セッションの平均時間は 84.8 秒（標準偏差: 46.6）であり、Double Cross Model の表現を含まない 183 セッションの平均時間は 89.6 秒（標準偏差: 56.8）であった。t 検定の結果、Double Cross Model の表現を含むセッションの時間が有意に短縮されていた。Double Cross Model の表現を含む 17 セッションのうち目的達成しなかったものは 1 セッション (5.8%) であり、Double Cross Model の表現がない 183 セッションのうち目的達成しなかったものは 18 セッション (9.8%) であった。これは、Double Cross Model の表現が位置の曖昧さを解消するのに有効であることを示唆している。

空間論理の分野では、Double Cross Model のフレームで示されるような 3 点の関係を表す言語表現が効率的である。本研究では実験協力者が実際に生成した言語表現を分析したが、実践的には移動を含む位置指定の課題における発話生成の研究において Double Cross Model がよ

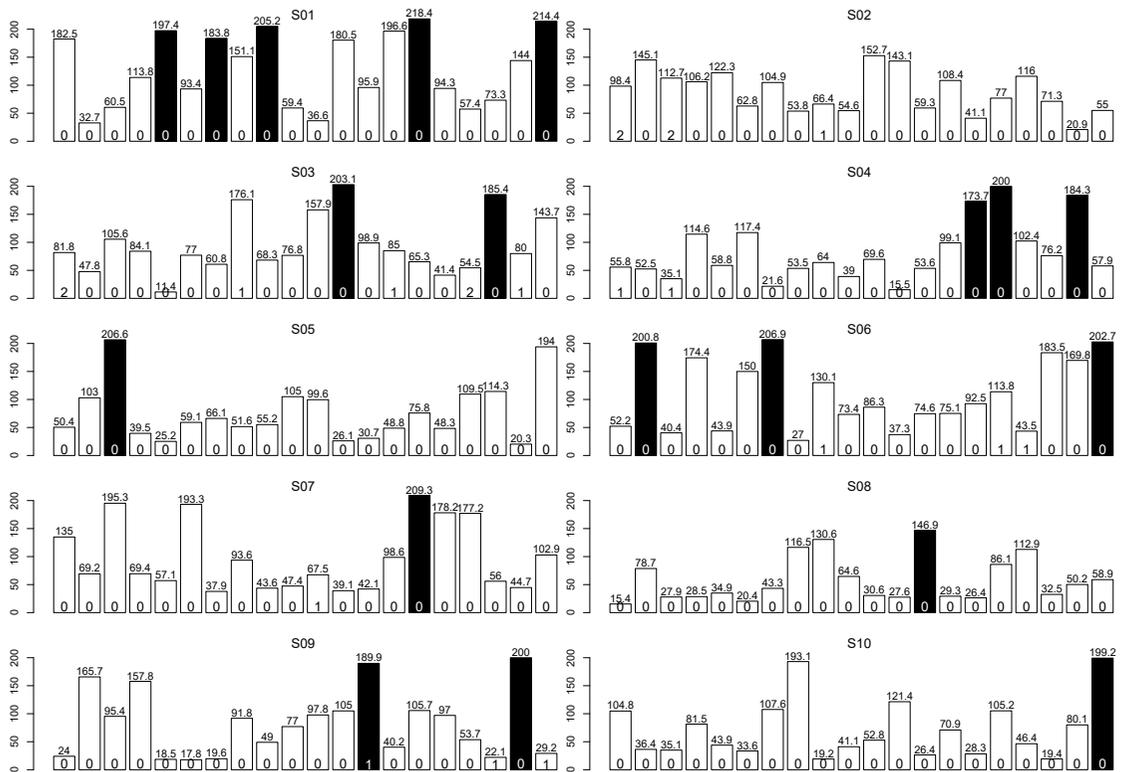


図 15 各セッションに要した時間 (実験協力者 10 人 × 20 セッション)

り重要だと考えている。例えば、タクシーの配車や自動運転など、詳細な移動指示を明確に表現する言語生成モデルに焦点を当てた研究において Double Cross Model が重要であろう。

5 おわりに

本研究では、空間参照表現のフレーム記述手法として Double Cross Model に基づく記述手法を提案した。従来の空間参照表現の記述は、主に同一参照、部分と全体の関係、および位相情報に基づく隣接関係に焦点を当てていた。実用的には、隣接しないランドマークや空間内実体に対して、相対位置情報を規定するにあたり、方向を表現するために 3 つ以上の点が必要である。3 点記述を実現するために Double Cross Model によるフレーム記述を提案した。記述手法においては、Double Cross Model における同型についても触れ、アノテーターがタグ付けしやすいものを使用することを提案した。さらに、離れたものに対しては、フレーム内に絶対的および相対的な距離を記述することを提案した。

東京都内（表参道と豊洲）で撮影した 360 度カメラ静止画を刺激として、位置情報を共有する発話収録実験を行い、10 人の参加者から音声データを収集した。収集した音声データに対して、RCC-8 に基づく位相情報のアノテーションと、Double Cross Model に基づく相対位置情報の記述を行った。アノテーションの結果から、同一参照、隣接、および部分と全体の関係に基づく言語表現が一般的であり、Double Cross Model を使用したフレームの記述の必要性は限定的であることがわかった。限定的ながらも Double Cross Model を用いて表現する必要がある相対位置情報表現が出現したセッションでは、位置情報の共有の達成率が高く、達成までの時間が短かったことがわかった。これは、相対位置情報表現は位置の曖昧さを解消する上で重要であることを示唆している。

空間論理の観点から見ると、位置情報の曖昧さを同一参照関係、隣接関係、および部分と全体の関係だけで解消することは非効率である。3 つの離れた対象の相対位置情報を使用する方が、2 つの対象の同一参照関係、隣接関係、および部分と全体の関係だけを使用するよりも、位置情報を曖昧さから解消するのに効率的である。この観点から、3 点に基づく相対位置情報表現の使用は、言語を介して位置情報を記述する際に重要である。近年、生成 AI の研究が盛んになっている中で、特に自動運転などのタスクにおいて、静止画、動画、地図情報に基づいて 3 点に基づく相対位置情報表現の生成が重要である。この研究は、この目的のための基礎的な研究となる。

今回、収集したデータは残念ながら 3 点に基づく相対位置情報表現の使用が限定的であった。本研究を発展させるために、位置情報参照表現(大村 他 2024)・経路情報参照表現(川端 他 2024)の大規模収集を進める。今後、位置情報参照表現・経路情報参照表現の大規模データに対してアノテーション仕様を検討し、提案手法で対応できない表現がないかについて確認を進める。

謝 辞

本研究はホンダ・リサーチ・インスティテュート・ジャパン-国立国語研究所共同研究プロジェクト「行き先目標物の参照表現に関する日本語話し言葉の分析」・国立国語研究所基幹型共同研究プロジェクト「アノテーションデータを用いた実証的計算心理言語学」・科研費 JP22K13108, JP19K13195 によるものです。

付録

A アノテーションデータの詳細

A.1 発話収録実験の仕様

以下では発話収録実験の概要について示す。

想定：実験参加者 2 名（A【位置情報伝達側】・ロボ役【位置情報受理側】と呼ぶ）による。実験参加者 A が周囲を見回して、位置同定に必要な視認できる情報をロボ役に伝達する。実験参加者 A はシステムの通常の利用者であるが、ロボ役はシステムであると想定する。

実験設定：実験参加者 A は一般募集し謝礼を支払う。ロボ役は実験委託先の方に担当してもらいシステムとしてふるまってもらう。実験参加者 A とロボ役は別々のブース／部屋にいてマイクとスピーカーを用いて音声のみをやりとりする。

実験参加者 A は、360 度カメラにより町中において撮像した静止画像を確認し、画像上に示された位置を同定可能な情報を自由に発話する。

ロボ役は、地図ソフトウェアを確認しながら実験参加者 A が見ている 360 度カメラ画像に指定された位置を聞き出す。このときロボ役は言えることが制限されており、できるだけ人のような受け答えを行わないように注意する。このためロボ役は実験中に言えることの例が書かれた「言えることリスト」を参照しながら実験参加者 A と対話する。

時間設定：1 実験で 1 つ有効な発話を収集できると想定する。1 実験当たり 2 分 30 秒間対話を行うこととし 1 実験参加者 A あたり 20 実験を行う。（各実験ごとに 30 秒間の休憩をし、さらに 10 実験後に 5 分の途中休憩を設定する。）1 人当たりの拘束時間は 1 時間 30 分以内を想定する。

実験参加者 A の想定と人数：日本語を母語とし不自由なく発話できモニター上の静止画像の認識ができ、マウスを使った簡単な画像ビューワーの操作ができること。男女同数とし 20 代から 50 代まで均等にする。

収集物：実験参加者 A の音声と操作画面、ロボ役の音声と操作画面をそれぞれ収集する。

A.2 転記の仕様

以下では転記の仕様の概要について示す。

位置情報伝達側と位置情報受理側との対話音声を聞いて、発話時間情報を取得し、発話内容を書き起こす。

基本方針として、フィラーや言いよどみも含めて、聞こえたとおりにすべて書き起こす。

分析対象である位置情報伝達側（実験参加者 A）の発話の区切りは 300 ミリ秒以上の無発話区間を区切りとする。分析対象でない位置情報受理側（実験参加者ロボ役）の発話の区切りは

300 ミリ秒以上の無音区間があっても、1回の発話でまとめる。

書き起こしは基本的に全角とし、句点「。」は付与しない。読点「、」はひらがなが連続する箇所などに読みやすくなる程度に付与する。「?」「!」は付与しない。

口語特有のいいまわしは、そのまま聞こえるように書き起こし、通常の記事に補正しない。語尾やフィラーなどで、長音が感じられた場合、長音記号「ー」を使用する。笑い声は書き起こさずに「笑」のように記述する。聞き取れない箇所は「@」のように記述する。言い間違い、言いよどみは聞こえる音に近い表記かつひらがなで書き起こす。店舗名、地名等の固有名詞の表記はカタカナで書き起こす。アルファベットの綴りを読み上げている場合はカタカナで書き起こす。

A.3 メンションのアノテーション

分析対象である位置情報伝達側（実験参加者 A）に対して、フィラー・応答語・要求語・場所（参照場所・目標場所）・位置関係（参照場所の位置関係・目標場所の位置関係）についてアノテーションした。このうち、場所と位置関係のメンションのみを分析対象とした。

参考文献

- Clementini, E., Di Felice, P., and Hernández, D. (1997). “Qualitative Representation of Positional Information.” *Artificial Intelligence*, **95** (2), pp. 317–356.
- Freksa, C. (1992). “Using Orientation Information for Qualitative Spatial Reasoning.” In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pp. 162–178. Springer.
- Hernández, D., Clementini, E., and Di Felice, P. (1995). “Qualitative Distances.” In Frank, A. U. and Kuhn, W. (Eds.), *Spatial Information Theory A Theoretical Basis for GIS*, pp. 45–57, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Islı, A. and Moratz, R. (1999). “Qualitative Spatial Representation and Reasoning: Algebraic Models for Relative Position.” Tech. rep..
- 川端良子, 大村舞, 小西光, 浅原正幸, 竹内誉羽 (2024). 地図を刺激に用いた経路情報参照表現の収集. 言語処理学会第 30 回年次大会発表論文集, pp. 1345–1350. [Y. Kawabata et al. (2024). Chizu o Shigeki ni Mochiita Keiro-joho-sansho-hyogen no Shushu. Proceedings of the 30th Annual Meeting of the Association for Natural Language Processing, pp. 1345–1350.].
- Ligozat, G. É. (1998). “Reasoning about Cardinal Directions.” *Journal of Visual Languages & Computing*, **9** (1), pp. 23–44.
- Mani, I., Hitzeman, J., Richer, J., Harris, D., Quimby, R., and Wellner, B. (2008). “SpatialML: Annotation Scheme, Corpora, and Tools.” In *Proceedings of the 6th International Confer-*

- ence on Language Resources and Evaluation (LREC'08), Marrakech, Morocco. European Language Resources Association (ELRA).
- Minsky, M. (1974). “A Framework for Representing Knowledge.” Tech. rep., Massachusetts Institute of Technology.
- 大村舞, 川端良子, 小西光, 浅原正幸, 竹内誉羽 (2024). 地図を刺激に用いた位置情報参照表現の収集. 第30回言語処理学会年次大会発表論文集, pp. 1469–1474. [M. Omura et al. (2024). Chizu o Shigeki ni Mochiita Ichi-joho-sansho-hyogen no Shushu, Proceedings of the 30th Annual Meeting of the Association for Natural Language Processing, pp. 1469–1474.].
- Pustejovsky, J. (2017). *ISO-Space: Annotating Static and Dynamic Spatial Information*, pp. 989–1024. Springer Netherlands, Dordrecht.
- Pustejovsky, J. and Yocum, Z. (2014). “Image Annotation with ISO-Space: Distinguishing Content from Structure.” In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14)*, pp. 426–431, Reykjavik, Iceland. European Language Resources Association (ELRA).
- Randell, D. A., Cui, Z., and Cohn, A. G. (1992). “A Spatial Logic Based on Regions and Connection.” In *Principles of Knowledge Representation and Reasoning: Proceedings of the 3rd International Conference*, pp. 165–176.
- Renz, J. and Nebel, B. (2007). “Qualitative Spatial Reasoning using Constraint Calculi.” In *Handbook of Spatial Logics*, pp. 161–215. Springer.
- Scivos, A. and Nebel, B. (2001). “Double-Crossing: Decidability and Computational Complexity of a Qualitative Calculus for Navigation.” In Montello, D. R. (Ed.), *Spatial Information Theory*, pp. 431–446, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Zimmermann, K. and Freksa, C. (1996). “Qualitative Spatial Reasoning using Orientation, Distance, and Path Knowledge.” *Applied Intelligence*, **6** (1), pp. 49–58.

略歴

川端 良子：国立国語研究所特任助教。

大村 舞：国立国語研究所プロジェクト非常勤研究員。

浅原 正幸：国立国語研究所・総合研究大学院大学教授。

竹内 誉羽：株式会社ホンダ・リサーチ・インスティテュート・ジャパン，
リンシパル・エンジニア。

(2024年1月31日 受付)

(2024年4月25日 再受付)

(2024年6月4日 採録)