

# インスタンス管理問題における identity の取り扱いに関する考察

遠藤 諭<sup>†</sup> 古崎 晃司<sup>†</sup> 溝口 理一郎<sup>†</sup>

<sup>†</sup>大阪大学 産業科学研究所 〒567-0047 茨木市美穂が丘 8-1

E-mail: <sup>†</sup>{endo, kozaki, miz}@ei.sanken.osaka-u.ac.jp

**あらまし** オントロジーに基づいた知識（インスタンスモデル）を記述・利用する際には、インスタンスの振る舞いを考察し、計算機上で適切に取扱うための管理が必要である。特にインスタンスが持つ identity は、そのオントロジーの定義内容により、インスタンスが持つ性質が異なる。本稿では、インスタンス管理問題において、取り扱いたい問題に応じて、インスタンスが持つ identity のどのような性質を考慮すべきかを議論する。そして、その考察結果に基づいて4種類の identity の導入を提案する。

**キーワード** identity, インスタンス, オントロジー

## 1. はじめに

近年、知識システムの基盤となるオントロジーは様々なシステムやモデルで応用されるようになり、その有用性への認識は広がりつつある。オントロジーは知識（インスタンスモデル）を記述するための規約を提供し、その規約に基づいて知識を記述することで、その一貫性や再利用性を保つことが重要な役割のひとつである。オントロジーの基礎理論は、より適切なオントロジーを構築するための指針を提供するものとして様々な研究が行われており、中でもロール概念の理論は重要な課題の一つとして活発な研究が行われている[1,2,3,4,6,8]。しかし、このようなオントロジー工学の研究の進展に伴い、オントロジーに関する理論は充実してきている一方、インスタンスモデルにオントロジーにおける定義内容をどのように反映するべきなのかは、明らかになっていない部分が多く、多くの研究の余地が残されている。特に、ロール概念やロールホルダーのインスタンスが持つ identity は、そのオントロジーの定義内容に応じて様々な性質があり、異なるタスクを考えたときに取り扱う identity にも違いが生じてくる。また、このようなインスタンスにおける identity の取り扱いは、ロール概念の性質からより厳密な考察が必要となってくる。このような背景のもと、本稿ではインスタンスが持つ identity の性質について考察し、4種類の identity を導入する。

以下、2章ではロール概念のインスタンスを適切に扱う際に identity に関する考察が必要である理由と、一般的に広く認知されている identity の分類を述べる、3章では、本稿で考察の対象とする identity の強さの違いを中心とした問題を通して identity の性質を議論し、4種類の identity を導入する。続く4章と5章で今後の検討課題と本論文の総括を述べる。

## 2. インスタンス管理問題における identity

### 2.1. ロール概念に関する identity

ロールホルダーの identity はロール概念とそのプレイヤーとなる概念の identity より合成して決定される。例えば、レストランというコンテキストで定義されるロール概念「客ロール」をプレイヤー「人間」が担ったときに具体化される客ロールホルダー「客」の identity は、客ロールと人間の identity より決定される。ここで例として、あるレストランの月間の客の数をカウントすることを考える。このとき問題の対象となるのは、山田さんや田中さんといった「特定された客（人間）」まで言及されているわけではなく、このレストランの「客になった人間の数」である。つまり、客になった人間が誰であるのかといった個々の客を特定するような identity までは取り扱う必要がなく、客の数をカウントできる程度の identity を考えればよい。一方、常連客を優遇するなどするために顧客管理を厳密に行いたい場合には、誰が（どの人間が）客となったかという、個々の客を特定するような identity を考える必要がある。このように、客ロールやロールホルダーのインスタンス管理を適切に行うためには、どのようなタスクを想定するかによって、複数の種類の identity を考える必要があることが分かる。また、日本においてある「人間」が「日本首相ロール」を担った「日本首相」（ロールホルダー）を考えたとき、日本首相は日本に一人しかいないことから、プレイヤーとなる人間が A 氏から F 氏に変わっても日本首相ロールの identity は変わらないと考えることが出来る。しかし、過去に人間が日本首相となった回数によって、現在の日本首相は第何代日本首相ということが言えることから、A 氏が担った日本首相ロールと F 氏が担った日本首相ロールの identity は異なるという捉え方も出来る。以上のように、ロール概念のインスタンスが持つ identity は、複数の種類を考えることができ、適切なインスタンス管理のためには、identity の種類と性質の考

察が必要となる。そこで、本稿ではインスタンスが持つ identity について、想定するタスクによってどのような identity を用いる必要があるのか議論を行い 4 種類の identity の導入を行う。

## 2.2. 一般的な identity の分類

一般的に知られている identity は、

- (1)人々の頭の中にある認識と、現実の世界との一致を議論する identity (例: Please identify yourself)
- (2)(複数の) インスタンスの同一性を議論する identity に大別される。

(1)の identity は、「あの中に遠藤さんはいますか?」「はい、いますこの人です」といった対話や、「さっきすれ違った人は誰ですか?」「遠藤さんです」といった対話における答えを判断する際に用いられる identity である。(2)の identity は更に **Synchronic identity** と **Diachronic identity** の 2 種類に分類される。

### Synchronic identity :

時間に依存せずにインスタンスが同じものであることを取り扱う identity

例えば、「10時から11時までの1時間と、11時から12時までの1時間は同じか?」といった問いに対して、時間量としては同じだが、時区間としては別として取り扱うことができる。また「宵の明星」と「明けの明星」は異なるラベルを持つが、金星という同じものであるということに対応する identity。

### Diachronic identity :

2つの時点での同一性を議論する identity

例えば、人間のインスタンスである太郎さんが5年前と現在で同一人物であるかといったことや、自転車の部品を交換した際において交換前後での自転車のインスタンスの同一性を議論する際には、Diachronic identity が用いられる。

本研究では、Diachronic identity を中心として、2つの時点でのインスタンスの同一性に関する考察と、identity の種類について考察する。

## 3. インスタンスが持つ identity に関する考察

2つの時点でのインスタンスの同一性を議論する際には、どの程度厳密に同一性を判断するかという「強さ」に応じて、identity の種類を考えることができる。

ここで identity の強さは、同一性の判断基準の厳密さ、換言すると「identity の変化のし易さ」を表す。例えば、にぎり寿司のインスタンスを考える。このにぎり寿司のシャリが1粒増えたとき、シャリの量が違うので増える前とは異なるにぎり寿司になったと捉えることもできるし、この程度の変化であれば同一の identity を持ったにぎり寿司であると捉えることもできる。また、

イカにぎりのインスタンスがあるとき、そのネタをイカからマグロに交換したとすると、そのにぎり寿司は、マグロにぎりといった別の identity をもったにぎり寿司に変化したと捉えることができる。このように、2つの時点での同一性を議論する際には、何種類かの identity を考えることができる。本研究では、identity の強さや性質に応じて、以下に述べる 4 種類の identity を導入した。

## 3.1. Identity for exactness

**Identity for exactness** (以下 **Iex** と表記する):

寸分変わらず同一であるという identity

例えば、にぎり寿司のインスタンスを考えるとき、シャリの米粒が1つでも増やされたりすると、そのにぎり寿司は増やす前後でシャリの量に違いが生じる為、この違いにより **Iex** は変化する。これにより、米粒の増加前後でのにぎり寿司のインスタンスは **Iex** の意味で異なることになる。

ここで、**Iex** の定義が示す「寸分変わらず同一である」ということには、時間変化に伴うインスタンスの変化も含まれる。例えば、にぎり寿司のインスタンスは時間が経つにつれて鮮度は落ちていくと考えられるため、それに伴い **Iex** は変化する。また一般的に、現実世界にある個物は時間に伴い分子レベルでの変化が起きていることから、全てのインスタンスが持つ **Iex** は時間の経過に伴い変化していると考えられるため **Iex** は保たれない。すなわち、**Iex** が保たれるものがあるとなれば、時間変化に依存しないものに限定される。

しかし現実問題を扱う際には、米粒が一粒増えたとしても同じにぎり寿司のインスタンスであると捉える場合が多いであろうし、人間のインスタンスである太郎さんについて考えると、5年前の太郎さんと現在の太郎さんは同一人物であるといった同一性を認識することができる。このようなことを適切に扱うには、Identity for exactness のように厳密な identity よりも、弱い identity を導入する必要がある。

## 3.2. Identity for essentiality

**Identity for essentiality** (**Ies**):

本質属性の値によって決定付けられている identity

例えば、にぎり寿司は構成しているネタの種類によって、イカにぎりやマグロにぎりといった種類が決まるので、にぎり寿司における本質属性はネタの種類と考えることができる。このとき、あるイカにぎりが持つ Identity for essentiality を **Ies1** とすると、このイカにぎりのインスタンスはネタが変わらない限り、例えシャリの量が増えたとしても **Ies1** を保ち続け、その間同じのイカにぎりであると扱われる。しかし、そのイカにぎりのネタをイカからマグロに変えると、にぎり寿司

の本質属性が変化することになるので、Ies1 は別の identity (Ies2 とする) へと変化する。これによって、ネタの交換前後のにぎり寿司は別のにぎり寿司であると扱われる。

なおにぎり寿司における Ies を決定付けている本質属性はネタの種類と認識しやすいのに対して、現実世界に存在するものの本質属性を決定することには一般に難しく、何を本質属性と捉えるかについては恣意性が生じる。例えば、自転車の本質属性は、「人が移動に使うことができる」という機能的側面から捉えることもできるし、「フレームと車輪とハンドルを部品として持つ」といた構成要素から捉えることもできる。自転車の本質属性を構成要素から捉えるとした場合であっても、必要な部品として「ハンドルやチェーン」を含めるか否かといった恣意性は残る。

このように、本質属性の決定には恣意性が認められるが、Ies ではその恣意性を問題としているのではなく、本質属性が決定しているときに、その本質属性に基づいてインスタンスの identity が決定されるという点を問題としている。すなわち、インスタンスの変化が本質属性の値の変化を伴っているか否かで、identity の変化を捉えることでできる。例えば、自転車の本質属性を「フレームと車輪」としたとき、自転車の部品を次々にはずしていく場合を考える。この際、部品が次々に交換されていく中で、どの時点で元の自転車ではなくなるのかといった問題が考えられるが、Identity for essentiality を用いると、この自転車は本質属性の値となっているフレームもしくは車輪が交換された時点で Ies が消滅し元の自転車でなくなると捉えることができる。

### 3.3. Identity for counting

インスタンスの数をカウントするというタスクにおいては、同じものを重複して数えると正確なカウントが達成されない。よってカウントする際に用いられる identity は同じものを重複して数えないという性質が重要となる。しかし、カウントを行うというタスクにおいては、3.1 前や 3.2 節で述べたようなインスタンスの詳細を考慮しなくとも、そのインスタンスがカウントの対象となっているかどうかだけわかればよい。また、ここで Ies を使ったとすると、数えている途中のにぎり寿司が本質属性が変化した場合、別の寿司になるため、もう一度その寿司を数えてしまうので、正しくカウントすることができなくなる。このように、カウントを行うというタスクで用いられる identity は属しているクラスに変化がない間は同じものであるということだけを主張することができればよい。以上の考察から、カウントするというタスクに用いる identity として identity for counting を導入する。

### Identity for counting (Ico) :

数を議論するための identity (Synchronic identity) 例えば、にぎり寿司が 10 個あり、その数をカウントするとき、にぎり寿司の詳細を考慮せずのにぎり寿司の数をカウントする際に、それぞれのにぎり寿司のインスタンスが持つ Ico (にぎり寿司) 1~10 をカウントすることによって、にぎり寿司が 10 個あることがわかる。さらに、この 10 個のにぎり寿司のなかで、イカにぎりは何個あるのかカウントする際には、イカにぎりクラスに属するインスタンスの Ico の数のみを数えればよい。

また Ico はインスタンスの数を主張する identity であることから、異なる二つの時点のインスタンスの数のみを比較したとき用いることができる。例えば、X 月 Y 日における駐輪場にとまっている自転車の数が 100 であったとすると、それら自転車の数は Ico の数をカウントすることによってわかる。また、この後の X 月 Z 日における駐輪場の自転車の数も同様に Ico をカウントすることによって 100 だとわかったとする。このとき、二つの日時における自転車の数は同じであるということがわかる。ただし、それらの二時点においてカウントした自転車のインスタンスが同一のものであるかどうかいうことは Ico から判断できない。カウントの対象としているインスタンスの詳細までを知ることができるのは、3.1 節および 3.2 節で述べた identity であり、Ico はそのようなインスタンスの内容を無視してよいときのみ用いることができる identity である。

### 3.4. identity for replacement

例として、マグロのにぎり寿司が 5 個あり、それらをカウントしている最中に、カウント済みのマグロのにぎり寿司同士の間でネタが交換されたという状況を想定する。このとき、カウントする identity を Ies として考えてみる。まず、マグロのにぎり寿司はそれぞれ Ies1 ~5 を持っている。この identity をカウントすることによってマグロにぎりの数をカウントすることができるが、カウントしている最中において、4 つめのマグロにぎりをカウントした時点で、2 つめに数えたマグロにぎり (Ies2) と 3 つめに数えたマグロにぎり (Ies3) のそれぞれのネタが交換されたとする。交換によってそれぞれのネタのマグロはシャリから離れることから、そのにぎり寿司は、マグロの切り身とシャリに分かれてしまい、マグロのにぎり寿司ではなくなってしまう。よって、それぞれが持っていた Ies2 と Ies3 は消滅する。その後、交換後にネタのマグロがそれぞれのシャリの上に乗ると、それぞれはマグロのにぎり寿司となるので、新たな Ies (Ies6 および Ies7 とする) を持つ。すると交換前にカウントしていた Ies2 および 3 と交換後

の Ies6, Ies7 はそれぞれ異なることから、交換後の Ies6 および Ies7 もカウントされる。よって、未カウントである Ies5 と交換後に生成された Ies6, Ies7 をカウントすることになり、実際にはマグロのにぎり寿司の数が 5 個しかないが、合計として 7 個あるとカウントされてしまう。

このように、カウントしている最中にインスタンスの部品交換などのような動的な変化が発生すると、交換後にできた新たな Ies をカウントするため、インスタンスの数を正確にカウントすることができなくなる。

以上のような問題を解決するための identity として identity for replacement を導入する。

#### Identity for replacement (Ire) :

全体物を固定して部品を交換するときに、交換される部品の種類と数の多さに無関係にそれが途中でいつの間にか別のものになってしまうと、そのものであり続ける identity。

例えば、先ほどのマグロのにぎりのネタが交換される例において、マグロのにぎりは Ire を持ち、カウントにおいて、マグロのにぎりは、それぞれ Ire1~5 を持つ。4 つめの Ire4 をカウントするときに、2 つめと 3 つめに数えたマグロのにぎりのネタ交換が行われとしても、それぞれが持つ Ire2,3 は保たれる。よって Ies のように新たな identity の生成が起きない為、残りの Ire5 だけのカウントし、マグロのにぎり寿司が 5 個あることが正確にカウントできる。このように、Ies では扱おうことができなかった部品交換が起きる可能性がある動的なモデルにおいて、Ire では正確なカウントを達成することができる。

#### 4. 検討

本稿で新たに導入した 4 種類の identity は、その強さは以下の順序に従う：

$$Iex > Ies > Ico = Ire$$

より深い考察が必要であると考えている。例えば、Ico であるが、カウンティングに時間が必要であると仮定しているが、そこに存在するインスタンスの数は存在する時点で確定しており、理論的にはカウンティングには時間は不要であることに注意が必要である。以上の内容を踏まえ、Ico = Ire に関しては今後十分な検討が必要である。

ロール概念のインスタンスにおいては、ロール概念の性質よりそのインスタンスはどの identity を持ち、扱うタスクにおいてはこういった identity の取り扱いが必要になるのか考察する必要がある。そして、ロール概念のインスタンスが持つ identity が異なるとき、その違いからロールホルダーの identity に与える影響も異なると考えられる。また、identity for replacement では、

全体物を固定して部品交換をすることができるが、このときに固定できる全体物がクラス定義において、どの上位概念まで固定可能な制約を与えるべきかなどが考えられる。また、このような identity の考察から得られるオントロジー構築の指針については、議論の余地が残されている。さらに、計算機上で取り扱う ID と identity の理論的対応など別途検討が必要である。

#### 5. まとめ

本稿では、identity に関する考察として、インスタンスが持つ本質に関わるタスクや、カウントにおけるタスクより、4 種類の identity の導入を行った。これらは、扱いたいタスクに応じてそれぞれ用いることによって、そのタスクを解決することができる。

今後はこれらの考察を深めていくと共に、ロール概念のインスタンスが持つ identity と、扱うタスクに応じてどのような identity を導入すべきかといった理論的枠組み構築と、計算機システム上への実装を進める。

#### 文 献

- [1] 砂川, 古崎, 來村, 溝口: コンテキスト依存性に基づくロール概念組織化の枠組み, 人工知能学会論文誌, Vol. 20, No.6, pp.461-472, 2005.
- [2] Steimann, F.: On the representation of roles in object-oriented and conceptual modelling, Journal of Data Knowledge Engineering, Vol.35, No.1, pp.83-106, 2000.
- [3] Masolo, C. et al.: Social Roles and their Descriptions, Proc. of the 9th International Conference on the Principles of Knowledge Representation and Reasoning (KR2004), pp. 267-277, 2004.
- [4] Loebe, F.: Abstract vs. Social Roles - A Refined Top-Level Ontological Analysis, Papers from the AAAI Fall Symposium Technical Report FS-05-08, pp.93-100, 2005.
- [5] 溝口理一郎, 池田満, 來村徳信: オントロジー工学基礎論, 人工知能学会誌, Vol. 14 (6), pp.1019-1032 (1999)
- [6] 古崎, 來村, 池田, 溝口: 「ロール」および「関係」に関する基礎的考察に基づくオントロジー記述環境の開発, 人工知能学会誌, vol.17, No.3, pp. 196-208, 2002.
- [7] 林雄介, 瀬田和久, 池田満, 金来, 角所収, 溝口理一郎: 概念間関係に関するオントロジー的考察～is-a, part-of, identity～, 電子情報通信学会技術研究報告, Vol. AI-98-40, pp.1-8 (1998)
- [8] A Model of Roles within an Ontology Development Tool: Hozo, Mizoguchi, R., Sunagawa, E., Kozaki, K., Kitamura, Y., J. of Applied Ontology, Vol.2, No.2, pp.159-179. Sep. 2007