

識別子とコード

Identifiers and codes

宮澤 彰
Akira MIYAZAWA

どんな研究？

データやプロトコルメッセージ等は、人間とコンピュータ、コンピュータ同士で情報を伝達する言語と考えることができる。これらの言語の中で識別子やコードの多くは、固有名詞に当たる機能を持つ。これらについて、自然言語との比較の中で考察する。

識別子とコード

- ここでいう識別子は、プログラミング言語の識別子ではなく、主としてネット上のさまざまな「もの」を識別する体系。たとえばDigital Object Identifierなど。
- ◆ 識別子は、固有名詞を持つような事物、コードは、より一般的にカテゴリー化されるような概念/事物に使われることが多いが、例えば国名コードのように固有名詞で表されるものにコードという名をつけることもある。また、バーコードのように基本的には有界でないものをコードと呼ぶこともある。
- ◆ 本質的には同一の性格を持つものであるが、主として日常的・継続的にメンテナンスを必要とするものを識別子、その必要のないものをコードと呼んでいると考える。

固有名詞

- 自然言語における固有名詞は、個人名、団体名、地名など、名詞のうちで個体を指し示すもの。ただし、たとえば、「源氏物語」は固有名詞か？という問題のように、どこまでが個体かという点のあいまいさは避けられない。
- ◆ 固有名詞には、普通名詞と異なるいくつかの性格がある。

固有名詞と識別子の対応

■ 類似している点

- ◆ 恣意的名前付けと社会的認知のシステム。例えば人名の場合、親(など)が名前をつける時点では全く恣意的であるが、戸籍への登録など社会的認知のためのシステムを通して、社会の中での機能を果たすようになっている。識別子やコードも、付与する時点では恣意的であるが、メタデータの流通など社会的認知のシステムを通して機能するようになる。逆に言えば、社会的認知のシステムを持たない識別子やコードは、社会的な機能を果たせない。
- ◆ 個別言語依存性の薄さ。普通名詞は意味が同じでも言語に依存して形が異なる(借用語はべつとして)。これに対し、固有名詞は音訳や同綴など、比較的単純な変換を通して異なる言語でも通用するようになる。識別子の場合、異なる種類のアプリケーションに対しても、比較的単純な変換で通用するようになっている。社会的に十分認知された例とは言い難いかもしれないが、RFC 3187 は、URN:ISBN:0-395-36341-1 のような形で、ISBNをURIとして使用することを提唱している。直接、固有名詞との対応は薄いですが、米国議会図書館(LC)は、
<http://id.loc.gov/authorities/sh85131823#concept>
のような形で、主題典拠ファイルレコードの識別子を、その主題に対するURIとして宣言している。
- ◆ 名前の禁忌。人名を直接呼ぶことに対する禁忌は、近代では薄れてきているものの、世界中に広く分布している習俗である。個人を識別する識別子もまた、社会的認知のある形では成立していない。名前の禁忌だけをこのことの原因にすることはできないが、現象としては類似している。

■ 相違している点

- ◆ 自然言語には、固有名詞の普通名詞化という現象がある。たとえば地名のシャンパンが、発泡ワインの一般名詞となるような例であるが、このような現象は、識別子・コードの世界には見られない。ただし、「.com 企業」のように、識別子が自然言語の中で普通名詞化した例はある。