

意味の理解とは何であるのか

—通信理論からのアプローチ—

What is ‘understanding meaning’?: An approach from the communication theory

山田隆弘*¹

Takahiro Yamada

*¹ 宇宙航空研究開発機構宇宙科学研究所

Japan Aerospace Exploration Agency/Institute of Space and Astronautical Science

This paper discusses the question, “What is understanding meaning?” In order to answer this question, we need a model that depicts how human beings process the meaning of language. This paper uses a model for communication called knowledge-based coding developed in the field of communication theory. This model was originally developed for image communication, but this paper generalizes it to cover human communication using language. The details of the generalized model are explained in the paper. Furthermore, this model is used to explain how human beings understand the meaning of sentences of natural language.

1. はじめに

言語の意味に関する理論的な研究は色々と存在するが [Cann 09, Kearns 2011 など]、人間が言語の意味を理解する過程に関して広く受け入れられている研究は少ないように思える。本稿では、「人間が言語の意味をどのように理解するのか」という問いへの回答を検討する。

人間が言語の意味を理解する過程を研究するためには、人間が意味を理解する過程を簡潔に表したモデルが有用である。本稿では、そのようなモデルとして通信理論で使用されているモデルを利用する。言語の意味の研究において通信理論が使用されるのは初めてではないが(過去に使用された例については 3.3 で述べる)、通信モデルに基づいて構築された言語理解のモデルが言語研究の中心的存在として使用された例はないように思える。言語の意味に関する様々な研究を相互に利用できるようにするには、どのような研究も共通のモデルに準拠する必要がある。言語の意味研究における共通のモデルを通信モデルに基づいて構築することが、本稿の第一の目的である。ここでは、特に知的符号化と呼ばれるモデルを採用する。知的符号化は、そもそもは画像通信のために考案されたものであるが、ここではそれを人間の言語を用いた通信に適するように一般化したものを使用する。本稿の二番目の目的は、このモデルを使用して「人間は言語の意味をどのように理解するのか」という問いに対する一般的な回答を探ることである。

以下、第2節から第4節において一般化された知的符号化モデルの詳細を説明する。第5節では、このモデルに基づいて意味の理解に関する仮説を提示する。第6節では、簡単な文についてその意味を理解するための条件を検討し、知的符号化のモデルが意味の理解の研究において演ずる役割を例示する。

2. 通信モデルと知的符号化

2.1 通信モデル

モデルに基づいた通信の理論的な研究は Shannon と Weaver の「通信の数学的理論」[Shannon 49]より始まったと考えられる。この理論に基づいて誤り訂正符号やデータ圧縮など

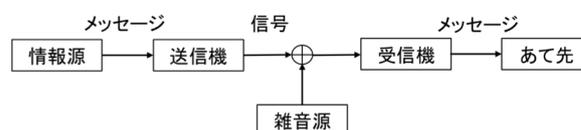


図1 Shannon-Weaverの通信モデル

様々な通信技術が開発されてきたのであるが、この理論は図1に示す通信モデルを基礎としている。このモデルは、通信システムを用いて情報源(情報の発信者)からあて先(情報の受信者)にメッセージ(情報)を送る方法を表現している。この図の送信機は、使用される通信媒体に適合した信号をメッセージから生成する機能を表す。通信媒体においては、生成された信号に雑音加わる。受信機では、雑音が付加された信号より元のメッセージを復元するための処理が行われ、復元されたメッセージは、あて先に届けられる。

機械的な通信には図1のモデルがうまく当てはまるが、人間どうしの言語を用いた通信を説明するためには、このモデルを少し修正の方が良い。なぜならば、人間どうしの通信では、(後述するように)人間の持つ知識が大きな役割を演ずるからである。

2.2 知的符号化

知識を用いた通信モデルとしては、[原島 88]の知的画像符号化が知られている。このモデルは、そもそもはテレビ電話などの画像通信におけるデータ圧縮技術への適用を目指して提案されたものである。しかし、このモデルは、画像だけでなく、人間どうしの言語を用いた通信にもうまく適用できる。本稿では、原島の知的画像符号化を一般化し、知的符号化を言語を用いた人間どうしの通信のモデルとして定義することにする。

一般化された知的符号化のモデルを図2に示す。このモデルでは、まず話者(speaker、情報の発信者)は何かの情報を聴者(hearer、情報の受信者)に伝えたいとする。そのために、話者は特定言語の文法と語彙を用いて発話(utterance)を行う。ここでは、伝えるべき情報は特定言語に依存しない情報であると考え、伝えるべき情報を特定言語の発話に変換するための装置を送信機と呼ぶことにする。聴者は、話者の発話を受けて元の情報を復元する。発話を元の情報に変換するための装置を受信機と

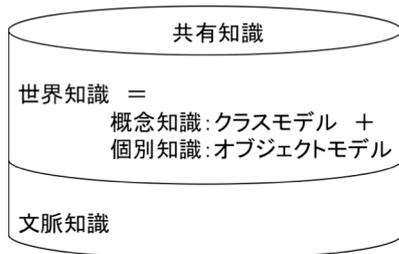


図3 共有知識の全体像

個々のものがどのような属性を具体的に持っているか、個々のもの間にどのような関係が存在しているかに関する知識である。このような知識をここでは個別知識と呼ぶことにする。

世界知識の表現形式としては、様々なものが研究されてきたが、最近ではオントロジーとして表現されることが多い[Gómez-Pérez 04]。また、オブジェクト指向モデリングの手法[Blaha 05]を用いてオントロジーを構築することもできる[Gasevic 10]。ここでは、知的符号化のモデルにおける共有知識は、オブジェクト指向モデリングの手法で表現したオントロジーであると想定することにする。オブジェクト指向においては、概念知識はクラスモデルで、個別知識はオブジェクトモデルで表現される。さらに、言語表現の意味構造もオブジェクト指向モデリングの手法で表現することが可能であり[山田 14]、ここでは発話の意味構造もオブジェクト指向モデリングの手法で表現されるものとする。図4には Unified Modeling Language (UML) [Booch 2005]のクラスモデルとして表した概念知識の例を、図5には UML のオブジェクトモデルとして表した個別知識の例を示す。

文脈知識の表現方法の検討は、今後の課題としたい。

5. 意味の理解とは何か

ここでは、前節までに提示してきた知的符号化のモデルに基づいて「人間は言語の意味をどのように理解するのか」という問いへの回答を探ることとする。

図2のモデルにおいて、聴者が発話を受信し、共有知識を用いて話者より伝えられた情報が正しく復元されたとする。これで聴者は発話の意味が理解できたとみなせるであろうか。

例えば、聴者は「ペドロ・マツアルはケツアルテナンゴに住んでいる」という発話より、<ペドロ・マツアルという人間の住所アトリビュートの値はケツアルテナンゴである>という意味構造を構築できたとする。しかし、聴者がペドロ・マツアルという人を全く知らず、ケツアルテナンゴがどこにあるのかも知らない場合、聴者はこの発話を理解したと言えるだろうか。一つの発話だけで話者が伝えたい情報を全て伝えることができるとは限らないので、もしかしたら「ペドロ・マツアルはケツアルテナンゴに住んでいる」という発話に続いて(同じ発話シーケンスの中で)ペドロ・マツアルと聴者との関係が明らかになる可能性もある。しかし、そうでない場合、聴者がこの発話を理解したとみなすことには無理があるであろう。「意味の理解」の定義にもよるが、ここでは、得た意味を自分の知識として使えるようになったときに「意味を理解した」と考える。自分が全く知らないことについての情報を得た場合、それを自分の知識として使うことは困難であるから、そのような意味は理解できないと考える。ただし、この結論については、もっと厳密な考察が必要であり、今後の検討課題としたい。以下では、上で述べたことを仮説として話を進める。(注:住所は、オブジェクトの持つアトリビュートとしてモデル化する場合と、オブジェクト間の関係(リンク)としてモデル化する場合とがあるが、ここでは前者であるとしている。)

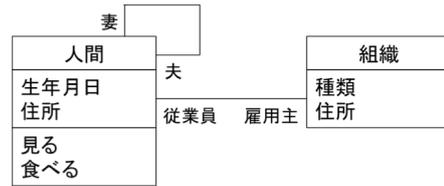


図4 UML のクラスモデルとして表した概念知識の例



図5 UML のオブジェクトモデルとして表した個別知識の例

上記の仮説をオブジェクトモデルに即して表現すると、聴者が発話を理解するための条件は「発話(あるいは発話シーケンス)から復元された情報の意味構造に現れるクラス、オブジェクト、アトリビュート値の中の少なくとも一つが聴者の保有している知識の中に存在していること」であると表される。先の発話に当てはめると、<ペドロ・マツアルという人の住所アトリビュートの値はケツアルテナンゴである>という意味構造に現れるオブジェクト(ペドロ・マツアル)とアトリビュート値(ケツアルテナンゴ)の少なくとも一つが既に保有している知識の中に存在している場合に聴者はこの発話の意味を理解したと言えることになる。ただし、上述のように、さらに厳密な検討が必要である。

上記のようにして理解された発話の意味は、聴者の知識に取り込まれ、それ以降の発話の処理に使用される。また、話者も自分の発話の意味が聴者によって理解されたと仮定し、その発話の意味を聴者が保有していると自分が推定している知識に取り込み、今後の発話で利用する。すなわち、発話の意味は共通知識に組み込まれることになる。共通知識は、オントロジーとして表現できるから、上記の仮説は「意味の理解とは、新たな知識を既存のオントロジーへ連続的に組み込むことである」とも表現できる。ここで、連続的とは「孤立した(他のクラス、オブジェクト、アトリビュート値との関係を持たない)クラスやオブジェクトやアトリビュート値が存在しない」という意味である。

上記のようにして発話の理解によって共通知識が拡充されるのであるが、既に存在している共通知識内の個々の知識も相互に連続的である必要がある。

6. ケーススタディ

ここでは、上記の仮説の内容を具体的に示すために、二つの文のタイプについて、人間が文の意味を理解するための条件を検討する。

6.1 x オブジェクトは C クラスに属する

まず、<x オブジェクトは C クラスに属する>という意味構造を持つ発話(文)を理解する条件を考える。具体的な発話の例として「ピタヤは果物である」を考える。この発話の意味構造は、<ピタヤオブジェクトは果物クラスに属する>と表される[山田

14]。(注:果物の一種であるピタヤはクラスなのか、オブジェクトなのかという議論もあり得るが、ここではオブジェクトであると考える。)

この発話を理解するための条件は何であろうか。聴者の知識の中にピタヤも果物も存在するが、「ピタヤが果物である」という知識が存在しなかった場合は、この発話を理解することができ、 \langle ピタヤオブジェクトは果物クラスに属する \rangle という意味が新たに聴者の知識に取り込まれる。

次に、聴者の知識の中にピタヤは存在するが、果物は存在しない場合はどうであろうか。この場合は、ピタヤが何であるかを理解したとは言えないが、「ピタヤは自分の知らない種類のものである」という部分的な理解はできる。逆に、聴者の知識の中に果物は存在するが、ピタヤは存在しない場合はどうであろうか。この場合も、発話を完全に理解したとは言いが、発話を「ピタヤという果物が存在する」と理解し、それを自分の知識に組み込むことは可能であろう。

最後に、聴者の知識の中にピタヤも果物も存在しない場合は、この発話は理解できないと考えるべきであろう。

上の議論をまとめると、 $\langle x$ オブジェクトは C クラスに属する \rangle という意味構造を持つ発話を理解する条件は、聴者の知識の中に x オブジェクトあるいは C クラスの少なくとも一方が存在していることであると結論付けられる。

6.2 x オブジェクトの w (リンク先)は y オブジェクトである

2番目に、 $\langle x$ オブジェクトの w (リンク先)は y オブジェクトである \rangle という意味構造を持つ発話を理解する条件を考える。具体的な発話の例として「ペドロ・マツアルの妻はマリア・サフビンである」を考える。この発話の意味構造は、 \langle ペドロ・マツアルという人間の妻はマリア・サフビンという人間である \rangle と表される[山田 14]。

この発話を理解するための条件は何であろうか。聴者の知識の中にペドロ・マツアルもマリア・サフビンも存在するが、「ペドロ・マツアルの妻はマリア・サフビンである」という知識が存在しなかった場合は、この発話を理解することができ、 \langle ペドロ・マツアルという人間の妻はマリア・サフビンという人間である \rangle という意味が新たに聴者の知識に取り込まれるであろう。

次に、聴者の知識の中にペドロ・マツアルは存在するが、マリア・サフビンは存在しない場合はどうであろうか。この場合、発話を「ペドロ・マツアルはマリア・サフビンという人と結婚している」と理解し、それを自分の知識に組み込むことは可能であろう。逆に、聴者の知識の中にペドロ・マツアルは存在しないが、マリア・サフビンは存在する場合も、発話を「マリア・サフビンはペドロ・マツアルという人と結婚している」と理解し、それを自分の知識に組み込むことは可能であろう。

最後に、聴者の知識の中にペドロ・マツアルもマリア・サフビンも存在しない場合は、この発話は理解できないと考えるべきであろう。

上の議論をまとめると、 $\langle x$ オブジェクトの w (リンク先)は y オブジェクトである \rangle という意味構造を持つ発話を理解する条件は、聴者の知識の中に x あるいは y が存在していること、および、 x と y が属するクラスが w というアソシエーション(同様のリンクの集合)を持つという知識が存在していることであると結論付けられる。

7. おわりに

本稿では、人間が言語の意味を理解するための条件を知的符号化のモデルとオブジェクト指向の知識表現に基づいて考察した。「発話を理解する条件は、発話の意味構造に現れるクラス、

オブジェクト、アトリビュート値の中の少なくとも一つが聴者の保有している知識の中に存在していること」であるという仮説を提示し、その仮説の妥当性をいくつかの例によって示した。現段階では、この仮説が十分に証明されたとは言いが、仮説のさらなる精密化も必要である。しかし、意味の理解に関する議論は、話者と聴者が知識を共有していることを前提とすべきであることの実例は示せたのではないかと考える。今後は、ここに示した仮説の精密化と妥当性の検証を行う予定である。

自然言語処理の分野では、オントロジーを利用して言語の解釈や推論を行うための研究が増えている[Nirenburg 04, Chimiano 14, 中村 15 など]。本研究も意味とオントロジーの双方の観点からさらに発展させていきたいと考えている。

参考文献

- [中村 15] 中村 絢子, 峯島 宏, 戸次 大介: オントロジーを用いた自然言語の推論に向けて, 言語処理学会第21回年次大会, pp. 309-312 (2015).
- [原島 88] 原島 博: 知的画像符号化と知的通信, テレビジョン学会誌, Vol. 42, No. 6, pp. 519-525 (1988).
- [山田 14] 山田 隆弘: オブジェクト指向に基づく意味構造の記述について, 2014 年度人工知能学会全国大会, 214-OS-08a-4 (2014).
- [Blaha 05] Blaha, M., Rumbaugh, J.: *Object-Oriented Modeling and Design with UML*, Second Edition, Prentice Hall, 2005.
- [Booch 2005] Booch, G., Rumbaugh, J., Jacobson, I.: *The Unified Modeling Language User Guide*, Second Edition, Addison-Wesley, 2005.
- [Cann 09] Cann, R., Kempson, R., Gregoromichelaki, E.: *Semantics*, Cambridge University Press (2009).
- [Chimiano 14] Chimiano, P., Unger, C., McCrae, J.: *Ontology-Based Interpretation of Natural Language*, Morgan and Claypool Publishers (2014).
- [Gasevic 10] Gasevic, D., Djuric, D., Devedzic, V.: *Model Driven Engineering and Ontology Development*, Springer-Verlag (2010).
- [Gómez-Pérez 04] Gómez-Pérez, A., Fernández-López, M., Corcho, O.: *Ontological Engineering, with examples from the areas of knowledge management, e-commerce and the semantic web*, Springer-Verlag (2004).
- [Grice 89] Grice, H.: *Studies in the Way of Words*, Harvard University Press (1989).
- [Huang 07] Huang, Y.: *Pragmatics*, Oxford University Press (2007).
- [Kearns 11] Kearns, K.: *Semantics*, Second Edition, Palgrave Macmillan (2011).
- [Levinson 00] Levinson, S.: *Presumptive Meanings*, MIT Press (2000).
- [Nirenburg 04] Nirenburg, S., Raskin, V.: *Ontological Semantics*, MIT Press (2004).
- [Shannon 49] Shannon, C., Weaver, W.: *The Mathematical Theory of Communication*, University of Illinois Press, 1949.