



Title	言語処理におけるデフォルト値
Author(s)	舟阪, 晃
Citation	大阪外大英米研究. 1994, 19, p. 71-83
Version Type	VoR
URL	https://hdl.handle.net/11094/99169
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

言語処理におけるディフォルト値

舟 阪 晃

1 まえがき

本稿の目的は、言語処理の過程にディフォルト値の存在を主張し、言語処理のいくつかのレベルでディフォルト値の実例を概観することである。詳述については、稿を改めることになる。

また、言語習得の過程におけるディフォルト値も、興味深いテーマであるが、本稿では扱っていない。

2 言語処理とディフォルト値

人間の頭脳が言語を処理しているのを観察していると、その物理的な容量の割には、よく働いているといえる。機械による言語処理が直面する多くの問題と対比してみると、頭脳の「偉業」に驚かされる。が、現在のところ、頭脳の働きについての実態はほとんどなにもわかっていないといってよいぐらいである。しかしながら、考えてみると、頭脳というのは、とりもなおさず、われわれ自身のことであり、われわれの頭脳は、自分自身のことをほとんど「知らない」ままに、実によく働いているという不思議な状況が見えてくる。

このところ、心理言語学関係の文献に触れる機会が多く、その観点から、特に、英語の処理を考えてみると、人間の頭脳は、分かり切っていることは、わざわざ計算しないで、「既定値」で処理をし、予想外の新しい事柄についてのみ注意の対象にしているのではないかと思わせるところがある。言い換えれば、すみからすみまで注意を払っているのではなく、手抜きが可能など

ころは、できるだけ手抜きをしようとしているふしがある。上記の「既定値」をデフォルト値とよぶ。

英文を読む場合、英語がよく分かっている人は、現在処理中の事柄を踏まえて、デフォルト値を使って、つぎに何が来るかを先読み (lookahead) する傾向がある。このことは、いわゆる袋小路文が存在することで確認できる。つまり、袋小路文は、デフォルト値に対する違反であるといえよう。

さらに、卑近な例をあげるならば、英語の書取を学生に課している場合、英語がよく分かっている学生は、文のすべてに注意を払っているのではなく、重要なポイントのみに注意を集中し、分かり切っているところは、デフォルト値で処理をしている様子が見ええる。逆に、英語が身についていない学生の場合は、文全体に注意を払い、そのため、処理に時間がかかり、書き取るための余裕がなくなることになる。書取の成績の悪い学生から対処法を質問され、「君は文のすみずみまで注意を向けすぎているのではないか」と聞き方によっては禅問答のような返答をするのは、この理由によるのである。

一般的にいて、言語教育の目標は、いかに早く、効率よく、学習者に、デフォルト値を身につけさせるかということになろう。

3 デフォルト値の定義

これまで、デフォルト値は、「既定値」としてきたのであるが、実は、この用語の意味がかならずしも明確であるとはいえないところがある。つぎに、先学の記述をあげ、検討を加えておく。

(1) A default value is a typical value, which may be assumed if there is no evidence to the contrary. (Garnham 85:120)

(2) The language user assumes that the unit being processed is its "normal" or "canonical" form unless the unit is overtly marked to the contrary. (Prideaux - Baker 86:32)

(3) A default role filler is a role filler that is usually, but not always, present in members of a class. (Reiter 90:262)

(4) 反証が提示されない限りは当然のこととされる対象の特徴：たとえば、鳥のデフォルト値は、二つの翼、羽毛、尾、さえずる能力など。(Johnson-Laird 88:250)

(5) Default values are those for a slot that are used if no specific contextual information is supplied, Default values define normal cases. But they can be overridden in nonnormal situations. (Lakoff 87:116)

以上の記述のうち、(1)、(2)では、デフォルト値は、反証がない限り仮定される典型的な、ふつうの値、単位と認められる。それに対して、(3)、(4)では、範疇と、それを構成する素性の両方に注意が向けられ、範疇の特徴を構成する素性のうち、いつでもではないが、通常、存在する素性がデフォルト値となる。

このことから、デフォルト値には二つの側面があることに注意する必要がある。一つは、範疇内の素性に注目する観点で、ある対象物は、範疇内の素性をどれだけ多く、また、素性のうち典型的なものをどの程度もっているかにより、つまり、素性の束によって、デフォルト値がきまることになる。どちらかといえば、因子分析により規定できる普遍的素性に関するもので、もし、素性が、範疇を、必要・十分条件をみたすように設定できるとすれば、有効に働くであろう。

他方は、素性による条件をみたした範疇に対して、言語使用者の認知的・社会的判断をも加えて判定したデフォルト値で、いわば、内的な素性の存在は当然のこととし、範疇の「全体像」を、より重視している立場といえる。

言語の処理は、文化、心理などのすべてが関与するという観点を取り、森羅万象を素性で規定することが、現時点では、困難であるということ、また、範疇と範疇の境目がかならずしも不連続であるとは限らないことなどを考えあわせると、素性に重点をおいた方法より、「全体像」を認める方法の方が現実的であるといえよう。「なんとなく——らしい」というような感覚を、不連続な、また、程度の差が表現できにくい素性の束で記述することには、

かなりの困難が予測される。

最後に、前記の(5)では、文脈(context)、場(situation)によりディフォルト値が変わることが指摘されている。これは、言語処理を論じる場合、非常に重要で、無視することができない問題であるが、本稿にはおさまりきらないほど大きいテーマである。

4 ディフォルト値のレベル

これまでの議論では、ディフォルト値は、名詞の範疇についてのみ言及されているように思われるが、言語現象のすべてのレベルについて、検討されねばならない問題である。

二人——仮にA、Bとして——の間に対話があるとき、Aの頭脳の中にあった事象が、言語(非言語的伝達手段も含めて)という媒体によって、Bの頭脳の中に「転送」される。頭脳の中の事象と、媒体の質的な違いを考えると、このような「転送」が可能になること自体、驚きに値するが、そのときにディフォルト値が大きな働きをしているのではないかというのが、わたくしの考え方である。

「行間を読む」という表現があるが、言語の発信者は、意識的、無意識的かは別として、伝えるべき内容のかなりの部分を、行間に託しており、受信者は、託された行間の情報を、いわば、自分の責任において補い、両者の対話が成立しているといえる。むしろ、言語が明示的に伝える事象は、ごく限られたものであろう。受信者側の、ディフォルト値による補足がうまくいけばいくほど、発信者の頭脳にあった事象が、「転送」されたように見える度合いが高くなるといえる。

ディフォルト値には、いろいろのレベルが認められる。提示の便宜上、言語外的なもの、言語内的なものとを区別し、さらに、必要な場合は、そのなかに、レベルを区別しながら、概観していこう。

4. 1 言語外的なレベル

言語の内部のレベルについては、これまでにふつうに言及されているので、問題は少ないが、言語外的という表現は、ひとにより解釈が異なる可能性が大きく、あまり立ち入ると、泥沼に足を取られることになると思われるので、ここでは単純化して、以下のようにとらえておく。

人間の間に対話がある場合、両者のコミュニケーションが破綻なく実現されるためには、同一の言語社会を共有するという形で、デフォルト値が設定されていなければならない。さらに、両者とも人間であるという点から同じ波長の概念作用を共有しているということもいえよう。素人的に言うならば、「共通の知識」、「共通の常識」、「共通の世界観」があるから、コミュニケーションが成立するということになる。

このような言語外的なレベルにかかわる事象についての先学の提案の中には、Minsky の frame、Schank - Abelson の script、Rumelhart の schemata などが含まれるが、いずれも、問題の性質上、かならずしも明確な形でモデル化されているとはいえない。が、遅から早かれ、人間の知識を表示するフォーマットが明らかにされねばならないという観点からすれば、いずれも、有益な試みであるといえる。

上記の提案のうち、frame と script は、比較的よく言及されるので、その説明の原文を二、三紹介しておく。

(6) A frame is a collection of slots and slot fillers that describe a stereotypical items. (Stillings, *et al.* 87:151)

(7) ... This frame represents a prototype of the kind of entity in the world that falls under the concept. (Garnham 87:39)

(8) (A script is) a representation of a frequently occurring sequence of events, usually of a social nature, such as going to a restaurant or going to a doctor. (Von Eckardt 93:194)

上の説明を読み、なかでも、slot、filler、event などの用語を目にすると、かつてのアメリカの言語学者の Pike が思い出され、感慨深いものがある。

る。Pike の『言語』のなかでは、フットボールの試合や教会の礼拝などが、行動主義的な観点から精密に分析されており、人間行動のパタンの理解まで、研究の領域を拡大していこうとする Pike の考え方には注目に値するものがあつた。が、当時は、Pike の構想は、その射程が大きすぎ、「科学性」を追求するばかりに偏狭になつたアメリカ構造主義のなかでは、異端視された。しかし、いまや、偉大な先学の考え方を再考できるほど言語学も「オトナ」になってきたといえよう。

4. 2 言語内的レベル

言語処理にかかわる言語内的なレベルでは、分析や提示の際に設定されるレベルほど明確に区別されるものはなく、いろいろのレベルが相互に干渉して処理が行われているといえる。

数年前、つぎのような言葉遊びが流行つたことがあつた。

(9) A : 「ミリン」と、10回繰り返してってください。

B : 「ミリン」「ミリン」……

A : 鼻が長い動物は？

B : キリン。

A : 残念でした。象でした。

文字で書くとほとんど問題が生じるとは思えないが、実際、口頭で試してみると、かなりの人がうっかり間違ってしまうということを経験的に知っている。

「ミリン」と「キリン」は、どのレベルで接点をみつけたのであろうか。「ミリン」という言葉を発したときに、「キリン」が想起されているのだろうか。「…が長い動物」という表現を聞いたときに、先行している「ミリン」が引き金になって「キリン」が想起されたのだろうか。そのとき、「鼻が…」という表現の部分はなぜ無視されたのか。いろいろ興味が引かれる問題が認められる。

言葉遊びというのは、言語処理の観点からすれば、馬鹿にできない興味あ

るテーマを提供してくれるが、ここでは、言語処理が、いろいろのレベルの情報を総動員して実施されているらしいということを指摘するだけにして、話題を戻して、文の解釈における文法的・意味的側面を中心に考察をすすめていきたい。

4. 2. 1 文解釈の最初の手がかりとしての「主語と動詞」

そもそも、言語による伝達の最終目標は、意味を伝達することで、言語使用を振り返ってみて、そのときの伝達内容は記憶しているが、文法形式についてはあまり記憶していないということから、自律的な文法論は、すくなくとも言語処理の観点からは、かならずしも示唆的な成果を約束するものではないと思われる。

英語の書き言葉の場合、発話は、左から右へ流れていくので、解釈も、基本的には、そのような方向で進めていかざるを得ない。

文全体の左は、通例、主語であり、解釈の最初の手がかりは、主語によって与えられるといえる。しかしながら、主語自体が与える情報は限られており、最初の解釈の単位は、主語と、通例つぎにつづく動詞との集合ではないかというのが、わたくしの考え方である。このことは、文の解釈に際して、いわゆる袋小路文が存在していることで証明できると思われる。つまり、主語だけで、後続する部分を先読みすることには限界があるが、主語と動詞の集合を解釈の出発点としてとらえると、それ以後の部分の構造は、かなりの確度をもって、予測できるので、その結果、場合によっては、袋小路文に誘導される可能性があることが、説明しやすくなる。

つぎの(10)、(11)の文を解釈するとき、(b)の方が、(a)より解釈の時間が短いという観察があるが、主語のつぎに動詞が来るというのが、英語の場合、一番期待されている、つまり、デフォルト値の構造であるということが確認できる。

(10) (a) [the girl] who [the boy] is pushing [t] is tall.

(b) [The boy] who [t] is pushing [the girl] is tall.

(11) (a) It is [the girl] who [the boy] is pushing [t] .

(b) It is [the boy] who [t] is pushing [the girl] .

(Osherson - Lasnik 90:185)

以上、統語論的な側面をみてきた。つぎの(12)は、(a) (b) とともに同じ単語が主語になっているが、後続する動詞によって、項のデフォルト値が決定される例である。

(12) (a) John wrote a letter.

(b) John received a letter.

主語の *John* だけでは、その項は特定できないが、動詞との集合としてみると、(12a) の *John* の項は、[agent]、(12b) のそれは、[patient] となり、それをもとに、後続する部分の解釈が予測される。なお、項の数や、その名称の妥当性については、ここでは立ち入らないことにする。

4. 2. 2 動詞と後続部分の解釈

前のセクションでは、「主語と動詞」の集合に焦点をあてたが、ここでは、動詞とその後続の部分の解釈に注目する。そのために、(13)のように、主語の位置に同じ単語が来る文をあげ、動詞の違いが後続する部分の処理にどのような影響を与えるかを考えてみたい。

(13) (a) The helicopter located the wreckage in the ravine.

(b) The helicopter discovered the wreckage in the ravine.

(Garrett 90:152-53)

処理時間は、(b) の方が長くなるという観察がある。動詞 *locate* は、目的語として、補文ではない名詞句を期待するが、期待通りの形になっているので、処理速度は相対的に早い。一方、動詞 *discover* の場合は、目的語として、補文を期待しているが、実際は、補文でない名詞句がきたために、処理速度が遅くなったというのが、Garrett の説明である。

「期待値」、つまり、デフォルト値がどのように決まるのかは、難しい問題であるが、文脈を別とすれば、「頻度」が重要な要因の一つであると考

えられる。ちなみに、手元にある Brown Corpus を調べてみると、*discover* の後に、補文でない名詞句がくる場合と、補文がくる場合の割合は、4 : 11 となっており、補文がくる場合の方が、3 倍ほど頻度が高い。

以上は、統語論のレベルで問題をとらえたのであるが、つぎに、項のレベルでの例をあげることにする。

(14) (a) John bought a book for Bill.

(b) John gave a book to Bill.

(c) John put a book on the desk.

いずれの場合も、*John*は動詞との関係で、項としては〔agent〕をとるといえるが、動詞の後にくる項は、動詞によって決まってくる。

(14 a) では、必須の項は〔theme〕で、〔beneficial〕は随意的な項である。したがって、*buy* は、*give*とよく並行的に扱われるが、実は、動詞としての結合価は2である。つぎに、*give*は、必須の項として、〔theme〕と〔dative〕を必要としており、結合価は3である。さらに、(14 c) の *put* は、同じく結合価は3であるが、必須の項は、〔theme〕と、〔loc〕または〔direction〕である。このように、どのような項が、いくつくるかは、動詞、または、主語+動詞の集合が、決めているといえる。

つぎに、項よりも下位のレベル、いわば、選択素性のレベルでのデフォルト値について、動詞を中心にして、触れておきたい。

(15) (a) John hit Harry.

(b) John kicked Harry.

(c) John struck Harry.

(15)の文中の動詞は、すべて、デフォルト値として、随意的項〔instrument〕をとると考えられる。しかしながら、項より下位の選択素性のレベルで考えてみると、〔instrument〕の下位のレベルでは、動詞ごとに、デフォルト値が決まっているといえる。たとえば、*hit* のデフォルト選択素性は<hand>、<bat>、*kick* のそれは<foot>、*strike* のそれは、<stick>、<bat>、などであろう。

ここにあげた動詞の場合は、デフォルト値は、すべての人に異論がないとはいえないが、比較的特定しやすい例である。

つぎに、文脈により特定される例をみてみよう。たとえば、(16 a) を読んだ被験者は、(16 b) という認識をしたという報告がある。

(16) (a) John was trying to fix the birdhouse. He was pounding the nail.

(b) John was using a hammer to fix the birdhouse.

(Garnham 87:62)

(16 a) において、*pound* が、随意的項として選択する〔instrument〕の選択素性が<hammer>であるのは、*nail* という単語の存在による。

さらに、*open* などを例として考えてみると、〔instrument〕という随意的項の存在は主張できるが、それ以下の細かいデフォルト値は、文脈により決まるといわざるを得ない。たとえば、*open* の目的語が、*the locked door*、*the box*、*the letter*、*one's eyes*、*one's mind* などの場合、〔instrument〕の選択素性は、それぞれ違うといえよう。

以上の動詞は、デフォルト値として、随意的な〔instrument〕をもち、英語使用者は、そのことを承知の上で英語を使用している。したがって、〔instrument〕が明示的に表現されていないときは、選択素性のレベルをも含めて、デフォルト値で処理することが期待されている。逆に言うと、〔instrument〕が明示的に表現されているときは、それはデフォルト値ではないということの証明になる。たとえば、前述の *kick* の場合、選択素性としての<foot>は、二本の足のうち、左右にかかわりなく、一本を使うことを意味し、同時に二本の足で *kick* することはないということを当然のこととし、英語使用者は英語を使用しているといえよう。したがって、万一、両足で *kick* することがある場合は、そのように明示的に表現する必要があるし、また、一方の足で *kick* する場合も、左右に言及する場合は、デフォルト値からはずれるので、(17)のように明示的に〔instrument〕を表現する必要がでてくる。

(17) John kicked Harry with his right foot.

4. 2. 3 名詞のデフォルト値

すでにのべたごとく、範疇のデフォルト値は、範疇の素性によって規定する方法では、おのずから限界があると思われる。つまり、素性による記述には、必要・十分条件をみたすようなかたちで素性が確定されている必要があるが、それは現状では期待できない。また、いろいろの範疇は、かならずしも、お互いに不連続な集合として規定できるとはいえないことが多く、素性による記述は、实际的でない。たとえば、よく引かれる例として、いろいろのゲームに共通の素性を見つけることは不可能で、結局、「家族的類似性」(family resemblance) のような概念を使ってまとめるのが現実的であろう。

このような観点を前提し、範疇に対する成員の典型値について考えてみたい。Malt & Smith (84:261) の調査によれば、英語国民の場合、「鳥」という範疇の成員間の典型値の順位は、度合いの高い方から、*robin*、*bluebird*、*seagull* などとなり、また、「果物」については、*apple*、*peach*、*pear* などとなるという。英語国民の頭の中にある名詞範疇のデフォルト値の一面が明らかにされたわけで興味深い。

日本人の場合は、大学生を対象に、わたくしが調査したところによると、「鳥」については、1位は、「はと」で、2位は、「つばめ」、3位は、「すずめ」となっている。また、「くだもの」は、「りんご」、「みかん」、「ぶどう」の順になっている。

「鳥」の範疇に関するデフォルト値は、英語国民と日本人では、大きな差を示している。一方、「果物」の範疇の場合は、両者の間の差は小さく、「りんご」は共通している。一般的にいて、デフォルト値として想起される成員は、言語により、民族により違うわけで、実際のコミュニケーションの場では、この事実を留意しておく必要がある。自分にとってのデフォルト値が、他人にとっては必ずしもそうでないという自覚をもつことは、案外むづかしいものである。

最後に、名詞の中に、項がデフォルト値として内包されている場合を考えてみる。

(18) The man fled from the city.

たとえば、(18)の *fled* を名詞形にした *flight* は、動詞のときにもっていた項関係を内包している。このことは、*flight of the city* は認められないが、*flight from the city* は認められるということで証明できよう。英語使用者は、名詞の場合も、いろいろのレベルのデフォルト値を、コミュニケーションの場で、活用しているといえる。

5 あとがき

言語処理におけるデフォルト値の存在を示唆するのが本稿の目的であった。デフォルト値には、いろいろのレベルが認められ、そのいくつかを例をあげながら検討してきた。さらに詳細な研究が必要とされる問題は多いが、なかでも、項構造や、文脈の問題は、今後の重要な研究のテーマになると思われる。

(93年10月25日)

文献一覧

Dale, Robert, Chris Mellish & Michael Zock (eds.) (90) : *Current Research in Natural Language Generation*, Academic Press.

Garnham, Alan (85) : *Psycholinguistics : Central Topics*, Methuen.

----- (87) : *Mental Models as Representations of Discourse and Text*, Ellis Horwood.

Garrett, Merrill F. (90) : "Sentence processing" in Osherson, Daniel N. & Howard Lasnik (eds.) (90).

Johnson - Laird, Philip N. (88) : *The Computer and the Mind : An Introduction to Cognitive Science*, Harvard University Press. 海保博之他訳『心のシミュレーション』新曜社。

Lakoff, George (87) : *Women, Fire, and Dangerous Things : What Categories Reveal about the Mind*, The University of Chicago Press.

Malt, B. C. & E. E. Smith (84) : "Correlated properties in natural categories," *Journal of Verbal Learning and Verbal Behavior* 23, 250-69.

Osherson, Daniel N. & Howard Lasnik (90) : *Language : An Invitation to Cognitive Science* : vol. 1, The MIT Press.

Prideaux, Gary D. & William J. Baker (86) : *Strategies and Structure*, John Benjamins.

Reiter, Ehud (90) : "Generating descriptions that exploit a user's domain knowledge," in Dale, Robert, Chris Mellish and Michael Zock (eds.)(90).

Stillings, Neil A. *et al.* (87) : *Cognitive Science : An Introduction*, The MIT Press.

Von Eckardt, Barbara (93) : *What Is Cognitive Science ?* The MIT Press.

