

効率的表現を介した解釈と生成

堂坂 浩二

ATR 自動翻訳電話研究所

日英間対話文翻訳において、断片文や省略文の解釈と生成のモデルを構築する際の基盤となる表現の枠組を提案する。言語表現の運ぶ情報内容を記述するために、文脈に依存して異なる情報内容を運ぶことができる効率的な表現を用いる。効率的な表現が、適切な環境に埋め込まれていることを保証することによって、文脈に依存する要因を排除した複雑な表現に頼ることなく、効率的な表現を用いて、断片文や省略文の情報内容を記述できる。文の解釈は、その文の情報内容を記述する表現をより効率度の低い表現に置き換える過程としてモデル化され、生成はその逆の過程である。また、多様な効率度の表現を介して翻訳プロセスを実施できる。

Interpretation and Generation through Efficient Representations

DOHSAKA Kohji

ATR Interpreting Telephony Research Laboratories

A descriptive framework is proposed in which fragmentary or elliptical sentences are interpreted and generated in the process of translation between English and Japanese. The framework aims at handling an efficient representation which can convey various information according to the surrounding context. By guaranteeing that an efficient representation is embedded in the appropriate circumstance, we take advantage of the efficient representation to characterize the informational content of a sentence without resorting to the complex representation in which contextual factors are eliminated. The process of interpreting a sentence can be modeled as replacing a representation with less efficient one as well as the generation process can be modeled as the converse process. Moreover, the translation process can be executed using a variety of representations specified to different degrees of efficiency.

1 はじめに

言語表現は、それが使用される環境に応じて、多様な情報内容を運ぶことができる。この意味において、言語表現は、効率的である。特に、対話においては、いわゆる断片文や省略文と呼ばれる効率的な言語表現が頻繁に用いられる。対話における効率的な言語表現の使用の例を見るために、次に示す対話 (d1), (d2) について考察しよう。これは、国際会議の参加申し込みに関する問合せ側 (Q) と事務局側 (S) との間の対話であり、問合せ側と事務局側の各発話ごとに、意図する情報を伝達するための日本語文と英語文の幾つかが示されている。

対話 (d1)

(u1) Q: どのように登録料をお支払いしましょうか?

(u2) How shall I pay the registration fee?

(u3) S: 銀行振り込みをお願いします。

(u4) 銀行振り込みで支払って頂けますか?

(u5) 登録料は銀行振り込みで支払って頂けますか?

(u6) Bank transfer, please.

(u7) Could you pay by bank transfer?

(u8) Could you pay the fee by bank transfer?

対話 (d2)

(u9) S: どのように登録料をお支払いになりますか?

(u10) How will you pay the registration fee?

(u11) Q: 銀行振り込みをお願いします。

(u12) 銀行振り込みで支払いたいのですが。

(u13) 登録料は銀行振り込みで支払いたいのですが。

(u14) Bank transfer, please.

(u15) I would like to pay by bank transfer.

(u16) I would like to pay the fee by bank transfer.

対話 (d1) では、事務局側 (S) は、問合せ側 (Q) の質問に対して、依頼型の発話内行為 [8] を遂行していると見なすことができる。一方、対話 (d2) では、問合せ側 (Q) は、事務局側 (S) の質問に対して、約束型の発話内行為を遂行している。そういった発話内行為を遂行するために、両対話において、日本語表現「銀行振り込みをお願いします」と英語表現「Bank transfer, please」が使用されている。これらの表現は、使用される環境に依存して、異なる発話内行為の型 (依頼または約束) をもち、また、依頼や約束の内容となる行為 (「登録料を支払う」) に関する情報を環境に依存して伝達できるという意味で、効率的な言語表現である。また対話 (d1) において、文 (u4), (u7) は、「支払い」の対象が文の中で明示されていないという点で、効率的な言語表現である。

以上に見たような効率的な言語表現を用いることによって、我々は、冗長でない円滑なコミュニケーションを実現することができる。それゆえ、対話を処理する自然言語処理システムが、対話者間の円滑なコミュニケーションを手助けすることができるためには、効率的な表現を扱うための枠組をもつことが必要となる。対話文翻訳システムは、そのような自然言語処理システムの一例である。そこで、本稿では、日英間対話文翻訳において、効率的な言語表現を解釈 / 生成するためのモデルの基盤となる表現の枠組を提案する。特に、日本語対話における断片文や省略文と呼ばれる効率的な言語表現の解釈と生成の問題を中心として、日英間対話文翻訳に適用することを考慮しながら、枠組を示していく。まず、続く第2章において、対話文翻訳において効率的な言語表現を処理する際の問題点について考察し、この枠組で採用するアプローチについて述べる。

2 問題点とアプローチ

2.1 生成に関する問題点

従来、断片文や省略文の解釈を行なうモデルが多く提案されてきた。その代表的なものとして、対話参加者のもつプランの下に、文を解釈するアプローチ [7] と、文のセンターに基づいて (ゼロ) 代名詞の指示対象を同定するアプローチ [5] がある。しかし、プランやセンターといった文脈情報のみに基づくアプローチは、断片文や省略文の生成の問題に直接適用することができない。まず、プランに基づく文の解釈手法の骨子は、対話参加者のもつプラン構造を推論し、文が含む情報と矛盾しないような表現をプラン構造の中に見い出して、その表現に文が含む構成要素を当てはめるというものである。この時、プランに含まれるのは、断片文や省略文に直接対応するような効率的な表現ではなく、文脈に依存する要因を排除した完全な表現である。このアプローチを文の解釈に適用する際には、文脈の中で本来適切ではない簡潔過ぎる表現でも、解釈できてしまう。言い替えると、ある情報内容を伝達するための断片文や省略文を生成するという問題にこのアプローチを適用すると、どの情報を生成文の中に明示しなくても適切かという判断ができない。このことを見るために、次の対話 (d3) を見てみよう。

対話 (d3)

(u17) Q: どのように登録料をお支払いしましょうか?

(u18) S: 銀行振り込みです。 (*)

(u19) 登録料を銀行振り込みで支払って頂けますか? (*)

この対話は、対話 (d1) の事務局側 (S) の発話を文 (u18), (u19) に置き換えたものである。いずれの文も、対話 (d1) での発話に比べて、発話時に設定されている文脈の下では不適切である。まず、文 (u18) は、依頼として解釈できないので、英語依頼文 (u6)~(u8) から翻訳される日本語文としては不適切である。また、文 (u19) は、

支払いの対象として、‘登録料’を繰り返している冗長な発話という意味で、この文脈では不適切である。一方、登録料’を主題化した文(u5)は、それほど不適切とは思えない。今、この対話例の分析をプランのみを用いて分析することを考える。この文脈で主要なプラン構造は質問-応答に関するものである。文(u18), (u19)は、質問文から予測される応答文として矛盾する情報を含んでおらず、また、質問文で要求されている情報(‘支払い’の方法)は、確実に含んでいるから、質問-応答に関するプラン構造の中では正当な発話として認められ、英語依頼文からの翻訳として生成することが許されることになる。

また、日本語談話における局所的な注視状態のモデルに基づいた解釈手法として、先に述べたセンターに基づくゼロ代名詞同定手法[5]がある。今、文のセンターは省略され易いと考えて、その手法を生成に適用したとしても、対話(d1)の事務局側(S)の発話において、‘支払う’の対象を省略できることは説明できない。なぜなら、このとき、‘支払い’の行為者(Q)のみがセンターとなりえるからである。

以上の問題が生ずるのは、従来の解釈中心のモデルでは、効率的な表現が文脈に依存する要因を排除したような完全な表現に対して強制的に対応付けられており、個々の省略文や断片文が適切に使用できる環境を分類することを怠っているからである。

2.2 翻訳に関する問題点

対話(d1)において、断片文(u3)と(u6)との間の翻訳を行なうことを考えよう。これらの断片文は、それぞれの断片文の情報内容が等価であることが文脈から判断できる場合には、複雑な表現を介せずに相互に翻訳することができる。また、日本語断片文(u3)と英語自動詞文(u7)との間の翻訳を考えると、適切な文脈の下では、‘支払う(pay)’の対象を明示した表現を介する必要はない。同じことが、文(u4)と(u7)との間の翻訳についても言える。このように、様々な表現のレベルで翻訳を行なうことができれば、原言語文の解釈を不必要に深く行なう必要がなくなり、また、翻訳を行なうために不必要なほど複雑な表現を介する必要もなくなる。

2.3 アプローチ

以上に述べた問題点を克服するためには、断片文や省略文の生成や翻訳のためのモデルにおいては、文の情報内容を記述するために効率的な表現を直接に取り扱い、その表現が適切に使用可能な環境を明確に分類する必要がある。さらに、多様なレベルの表現を介した翻訳を実現するためには、同一の情報内容を記述するために効率度の異なる表現を用意するといったアプローチを採用する必要がある。

一方、言語の効率性を中心に据えた意味論として、状況意味論がある[1, Chapter 6 and 7]。そこでは、効率的な表現を直接に扱うために、効率的な表現の情報内容を表現とそれが埋め込まれている環境との組で記述する。すなわち、表現の情報内容の一部を環境に担わせるのである。続く第3章では、この状況意味論の考え方と状況理論の言語に基づいて、対話例を援用しながら、効率的な言語表現の解釈/生成/翻訳のための表現の枠組を示していく。さらに、第4章で、関連研究について概観し、最後に今後の課題を述べて結ぶ。

3 枠組

3.1 概観

まず、枠組の概観を述べる。この枠組では、文の情報内容を記述するために、効率的な表現を用いる。効率的な表現は、それが埋め込まれている環境に依存して異なる情報を運ぶことができる。また、同じ情報内容を記述するために、効率度の異なる多様な表現を用意する。環境への依存度が高い表現ほど効率度が高い。この考え方を実現するために、ある情報内容を運ぶために表現が埋め込まれる必要のある環境の型を分類する。環境は、表現が運ぶ情報内容の一部を担っているという意味で、表現の情報内容を与えるための資源として機能する。表現は、状況理論における事態として記述し、環境は、その環境がどういった状況の型によって分類されるかという観点から記述する。状況の型は、その状況が支持する事態とその事態を構成している対象物とから構成される。

対話の流れの中で既に導入された情報内容が、それ以降の文の使用に対して資源となる環境を与える。例えば、対話(d1)では、問合せ者(Q)の質問文の情報内容の一部が、続く事務局側(S)の応答文に対する環境となっている。この環境を利用することによって、効率的な表現を、完全な表現に直接対応付けることなしに、省略文や断片文の解釈と生成、翻訳を行なう。

なお、文を使用する際、それまでに導入された情報内容のどの部分が環境として利用されるかは、対話の構造に依存して決定される。対話の構造は、個々の発話間の関係や発話のグループの間の関係から構成される。対話(d1)では、質問と応答の対という発話間の関係が形成されており、その関係を通して、質問文の情報内容の一部が、応答文の環境として与えられている。発話間の関係としては、2つの発話が単に隣接しているという関係も考えられるし、発話者がもつプランによって支えられる関係も考えられる。しかし、本稿では、対話の構造を通して発話の

環境を得るための方法論については触れない。本稿での主たる興味は、個々の効率的な表現が使用できる環境を分類し、その環境を利用することによって、効率的な表現を介した解釈 / 生成 / 翻訳を行うことである。

3.2 表現

ここでは、対話 (d1),(d2) で現れた断片文や省略文の情報内容を記述するために用いる効率的な表現を幾つか示す。この枠組みでは、文の情報内容を、効率度の異なる表現で記述する。まず、文の統語構造を反映した文の型に対応する表現を考える。例えば、文 (u3)((u11)), (u6)((u14)), (u4), (u7), (u12), (u15) の文の型に対応する表現は、それぞれ次のような事態として書かれる。ただし、ここでは文 (u3) における「で」を格助詞としては扱わない。なお、 $[*s \mid *s \models \sigma]$ は、事態もしくは対象 σ を支持する状況の型を表し、「*」で始まる記号はパラメタを表す。

(r1) $\ll$ お願いする文, 銀行振り込み ;1 $\gg$.

(r2) $\ll$ Please 文, $[*s \mid *s \models \text{'bank transfer'}]$;1 $\gg$.

(r3) $\ll$ 日本語依頼型文, $[*s \mid *s \models \ll$ 支払う, { MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

(r4) $\ll$ 日本語約束型文, $[*s \mid *s \models \ll$ 支払う, { MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

(r5) $\ll$ 英語依頼型文, $[*s \mid *s \models \ll$ Pay, { MANNER $\rightarrow$ 'bank transfer' } ;1 $\gg$] ;1 $\gg$.

(r6) $\ll$ 英語約束型文, $[*s \mid *s \models \ll$ Pay, { MANNER $\rightarrow$ 'bank transfer' } ;1 $\gg$] ;1 $\gg$.

次に、文が遂行する発語内行為を反映した表現について考える。対話 (d1) では、応答文の情報内容を依頼型 (Requesting) の発語内行為を示す表現で記述し、一方、対話 (d2) においては、約束型 (Promising) の表現で記述できる。例えば、対話 (d1) の日本語応答文の情報内容は、次の表現 (r7) で記述され、対話 (d2) の日本語応答文の情報内容は、表現 (r8) で表される。また、同様の表現を用いて英語応答文の情報内容も記述できる。

(r7) $\ll$ Requesting, $[*s \mid *s \models \ll$ 支払う, { MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

(r8) $\ll$ Promising, $[*s \mid *s \models \ll$ 支払う, { MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

表現 (r1) の情報内容は、環境に含まれていた情報を表現の中に担わせることによって、表現 (r7) や (r8) を用いても記述できる。しかし、断片文 (u3) や (u11) の発語内行為をそれぞれ表現 (r7) と (r8) で直接記述する必要はない。すなわち、各対話 (d1), (d2) で設定される環境の下で共通に遂行される発語内行為の型を考えることができる。これを効率的願望型 (Efficient_Desiring) と呼ぼう。この型の行為は、話し手がある対象 (銀行振り込み) に関わる願望を持つことを伝達するために用いられ、環境に応じて、依頼型とも約束型とも解釈されるのである。この効率的願望型に対応する表現を次に示す。

(r9) $\ll$ Efficient_Desiring, $[*s \mid *s \models$ 銀行振り込み] ;1 $\gg$.

表現 (r8) の中で使われている状況の型は、断片文「お願いします」に含まれる名詞表現に対応する対象 (銀行振り込み) を構成要素としてもつような状況を示す。効率的願望の内容となる行為の型に関する情報は、この表現が使用されるとき環境が担っている。効率的願望型の表現 (r8) の情報内容は、環境に依存して、依頼型の表現 (r7) と約束型の (r8) とも記述できるという意味で、それら2つの表現よりも効率的である。また、表現 (r7), (r8) は、支払いの対象 (登録料) を明示していない表現であるという意味で、それを明示した次の表現 (r10), (r11) よりも効率的である。

(r10) $\ll$ Requesting, $[*s \mid *s \models \ll$ 支払う, { OBJE $\rightarrow$ 登録料, MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

(r11) $\ll$ Promising, $[*s \mid *s \models \ll$ 支払う, { OBJE $\rightarrow$ 登録料, MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$] ;1 $\gg$.

3.3 環境の分類

この章では、対話 (d1),(d2) の分析を中心に、第3.2章で示した効率的な表現が、省略文や断片文の運ぶ情報内容を記述するために、埋め込まれる必要のある環境を分類する。

3.3.1 効率願望型の表現を使用するための環境

日本語断片文 (u3)((u11)) と英語断片文 (u6)((u14)) が、効率的な願望型として分類できる情報内容をもつためには、次のような条件を満たす状況の型で分類される状況が、環境として存在している必要がある。

- (1) その状況を行為として見なすことができ、断片文に含まれる名詞表現に対応する対象が、その行為としての状況の中に矛盾なく取り込める。また、聞き手が、その行為の遂行に参加可能である。

対話 (d1) においては、日本語応答文が使用される時の環境としては、質問-応答の発話間の関係から、次の状況型で分類される状況が与えられている。英語側の環境も同様に与えられている。

(r12) $[*s \mid *s \models \ll$ 支払う, { AGEN $\rightarrow$ Q, OBJE $\rightarrow$ 登録料 } ;1 $\gg$ $\wedge$ 登録料 $\wedge$ Q] .

上の状況の型 (r12) で分類される環境は、日本語断片文 (u3) が効率的願望型の発語内行為として見なせるための条件 (1) を満足している。すなわち、この枠組では、状況が行為として見なせるとは、その状況が行為型の事態 σ を支持することを言う。状況の型 (r12) における関係「支払う」は、行為型である。また、対象が状況の中に矛盾なく取り込めるとは、その対象が、事態 σ のある引数ロールの値となり得ることを言う。対象「銀行振り込み」は、「支払い」の手段を表す引数ロール (MANNER) の値として適切である。また、聞き手が参加可能な行為とは、その行為を遂行することが聞き手とは無関係には行なえず、聞き手の協力や介入が仮定されているような行為である。国際会議の参加に関わる手続きとしては、申し込み者が登録料を特定の方法で支払うことは、事務局の了解が必要であるという共通の認識がある。

こういった環境に埋め込まれている時、表現 (r1) は、表現 (r9) と同様の情報内容をもつのである。さらに、効率的願望型の表現 (r9) が、依頼型 (r7) や約束型の表現 (r8) と同じ情報内容を運ぶためには、それぞれ次の条件 (2), (3) が必要である。

(2) 環境条件 (1) で述べた行為の行為者が聞き手である。(依頼型)

(3) 環境条件 (1) で述べた行為の行為者が話し手である。(約束型)

ここで重要なことは、効率的な表現 (r9) が以上のような適切な環境に埋め込まれているなら、同じ情報内容を、より完全な表現 (r7), (r8) に頼ることなく、簡潔な表現 (r9) を使って記述できるということである。

また、第 2.1 章では、対話 (d3) の中で断片文 (u18) が、依頼の英語文に対する日本語翻訳としては不適切であることを述べた。文 (u18) のような「ダ-パターン」に対応する表現は効率度が高過ぎて、対話例の中の環境を利用したとしても、効率的願望型の表現と同様の情報を運び得ないのである。一般に、「ダ-パターン」は、依頼もしくは約束としては使用できないように思える。例えば、次の対話 (d4) と (d5) を見てみよう。

対話 (d4)

(u20) Q: 用紙を送って頂けますか?

(u21) S: はい、そうです。(*)

(u22) はい、お送りします。

対話 (d5)

(u23) Q: 用紙を送って頂くことができるのですか?

(u24) S: はい、そうです。

(u25) はい、できます。

対話 (d4) では、問合せ者 (Q) から行為の依頼をされた事務局 (S) が、その依頼に同意し、用紙を送ることを約束している。しかし、その約束型の行為を、「ダ-パターン」では遂行できない。これに対して、対話 (d5) では、「ダ-パターン」は、約束ではなく、単に、能力に関する質問に対する応答として用いられているため、適切に使用できる。

3.3.2 構成要素を明示しない表現を使用するための環境

第 3.2 章で示した表現の幾つかは、その情報内容の構成要素のすべてを明示していないという意味で効率的であった。例えば、表現 (r7), (r8) では、「支払う (Pay)」の対象 (OBJE) が明示されていない。これらの効率的な表現は、適切な環境に埋め込まれていれば、より完全な表現 (r10), (r11) と同等の情報内容を運ぶことができる。こういった適切な環境に関する条件として次のものを考えることができる。

(4) ある引数ロールの値を明示していない表現が、その引数ロールの値を含む情報内容を運ぶためには、資源となる環境にその値として適切な対象が含まれている必要がある。さらに、引数ロールの適切な値として、その対象と区別できない同等の別の対象は、含まれていない。

条件 (4) を満たす環境に埋め込まれているなら、表現 (r7), (r8) は、対象ロールの値を明示しなくとも、その値を含む情報内容を運び得る。対話 (d1), (d2) での環境はこの条件を満たす。また、引数ロールの適切な値として、環境が含む対象を順序付けることは、センターの連続性に関する順序付け [5] などを利用することによって行なわれる。しかし、センターの連続性による順序付けを行なうだけでは、効率的な表現を使用できる環境を分類するには不十分であって、さらに、そういった順序付けによって、限られた環境の中で、適切な値となる対象が唯一決まることが必要である。次の対話 (d6) を見てみよう。

対話 (d6)

(u26) S: どのように登録用紙をお送り致しましょうか?

(u27) Q: ファクシミリでも郵便でも結構です。

(u28) S: それでは、ファクシミリでお送り致します。

(u29) それでは、お送り致します。(*)

対話 (d6) では、文 (u29) は、文 (u28) と同様の情報内容を伝達するためには、使用できない。これは、この文を使用する際の環境に、送る方法となる 2 つの同等な対象が存在しているためである。

また、文 (u4), (u12) の情報内容は、「支払う (Pay)」の行為者 (AGEN) が明示されていない表現 (r7), (r8) で記述できる。これらの表現が、「支払う (Pay)」の行為者 (AGEN) を含む情報内容を運ぶことは、待遇関係にや話

し手の共感度 [6] に関する環境を考えることによって説明できる。例えば、文 (u4) では、待遇表現 ‘支払って頂く’ が使われている。この待遇表現の使用によって、‘支払う’ の行為者として取り得る値に対して、環境からの制限が課せられる。すなわち、その値は、環境において、話し手より上位に待遇されるとわかる対象でなければならない。話し手と聞き手しか存在しない対話では、それは聞き手である。対話 (d1) の文 (u4) の使用の際には、それは、質問者 (Q) になって、‘支払う’ の行為者が質問者 (Q) であることがわかる [2]。また、そういった環境に埋め込まれることによって、表現 (r7) は、‘支払う’ の行為者ロール (AGEN) の値を明示しなくとも、その値を含む情報内容を伝達できる。

3.4 環境を利用した解釈と生成

第 3.3 章で述べたように、表現がある情報内容を運ぶためには、その表現が適切な環境に埋め込まれている必要があり、2 つの表現が限られた環境の中で使用されるなら、それらの異なる表現を用いて同一の情報内容を記述できる。このことを次の制約を用いて表現する。なお、 $\vec{x}$ は、パラメタのベクトルである。

$$(r13) R_1(\vec{x}) \Leftrightarrow R_2(\vec{x}) \mid C(\vec{x}).$$

すなわち、2 つの表現 $R_1(\vec{x})$, $R_2(\vec{x})$ が、タイプ $C(\vec{x})$ で分類される環境に埋め込まれているなら、それらの表現は、同一の情報内容を記述する。また、2 つの表現は、効率度が異なっていてよい。

この制約に基づいて、断片文の解釈と生成のプロセスを実現することができる。文の解釈は、まず、その文の情報内容を記述するための表現として、統語構造を反映した最も効率的な表現を得ることから始まる。その表現から出発して、その表現が埋め込まれている環境を利用しながら、文の情報内容を記述するためのより効率度の小さな表現を得ていく過程である。すなわち、環境が担っていた情報を表現自体に担わせるのである。この過程は、効率度の大きな表現から効率度の小さな表現を得る方向に制約を適用することによって実現できる。生成は、その逆のプロセスである。文の情報内容を記述する表現と環境が与えられた時、制約を解釈時とは逆向きに適用して、効率度の高い表現を得ていき、統語構造を反映した表現を得るまでこれを続ける。

また、この枠組では、情報内容の記述に様々な効率度の表現を用いることができるから、ある表現を解釈する際に、文脈に依存する要因を排除した完全な表現を常に用いる必要はないという利点がある。次に、第 3.3 章で述べた環境条件についての考察に基づいて、対話 (d1), (d2) の応答文を扱うための制約を示す。

$$(r14) \ll \text{お願いする文}, [s \mid s \models j] ; 1 \gg \Leftrightarrow \ll \text{Efficient_Desiring}, [s \mid s \models j] ; 1 \gg \mid [s \mid s \models \sigma], \\ \text{ただし、} \text{ReIn}(*R, * \sigma) \& \text{Action}(*R) \& \text{Appropriate-Value}(*R, * \text{role}, *j) \& \\ \text{Take-Part-in}(*hr \mid \ll \text{Hearer}, *hr ; 1 \gg, * \sigma).$$

$$(r15) \ll \text{Please 文}, [s \mid s \models e] ; 1 \gg \Leftrightarrow \ll \text{Efficient_Desiring}, [s \mid s \models e] ; 1 \gg \mid [s \mid s \models \sigma], \\ \text{ただし、} \text{ReIn}(*R, * \sigma) \& \text{Action}(*R) \& \text{Appropriate-Value}(*R, * \text{role}, *e) \& \\ \text{Take-Part-in}(*hr \mid \ll \text{Hearer}, *hr ; 1 \gg, * \sigma).$$

$$(r16) \ll \text{Efficient_Desiring}, [s \mid s \models *x] ; 1 \gg \Leftrightarrow \\ \ll \text{Requesting}, [s \mid s \models \ll *R, \{ * \text{role} \rightarrow *x \} ; 1 \gg] ; 1 \gg \mid [s \mid s \models \sigma], \\ \text{ただし、} \text{ReIn}(*R, * \sigma) \& \text{Action}(*R) \& \text{Appropriate-Value}(*R, * \text{role}, *x) \& \\ \text{Take-Part-in}(*hr \mid \ll \text{Hearer}, *hr ; 1 \gg, * \sigma) \& \text{Agent}(*hr, * \sigma).$$

$$(r17) \ll \text{Efficient_Desiring}, [s \mid s \models *x] ; 1 \gg \Leftrightarrow \\ \ll \text{Promising}, [s \mid s \models \ll *R, \{ * \text{role} \rightarrow *x \} ; 1 \gg] ; 1 \gg \mid [s \mid s \models \sigma], \\ \text{ただし、} \text{ReIn}(*R, * \sigma) \& \text{Action}(*R) \& \text{Appropriate-Value}(*R, * \text{role}, *x) \& \\ \text{Take-Part-in}(*hr \mid \ll \text{Hearer}, *hr ; 1 \gg, * \sigma) \& \text{Agent}(*sp \mid \ll \text{Speaker}, *sp ; 1 \gg, * \sigma).$$

$$(r18) \ll \text{Requesting}, [s \mid s \models \ll *R, *A ; 1 \gg] ; 1 \gg \Leftrightarrow \\ \ll \text{Requesting}, [s \mid s \models \ll *R, *A \cup \{ * \text{role} \rightarrow *x \} ; 1 \gg] ; 1 \gg \mid \text{Unique}(*R, * \text{role}, *x).$$

$$(r19) \ll \text{Promising}, [s \mid s \models \ll *R, *A ; 1 \gg] ; 1 \gg \Leftrightarrow \\ \ll \text{Promising}, [s \mid s \models \ll *R, *A \cup \{ * \text{role} \rightarrow *x \} ; 1 \gg] ; 1 \gg \mid \text{Unique}(*R, * \text{role}, *x).$$

$$(r20) \ll \text{英語依頼型文}, [s \mid s \models \sigma] ; 1 \gg \Leftrightarrow \\ \ll \text{Requesting}, [s \mid s \models \sigma] ; 1 \gg \mid \text{Requesting に関する準備条件などを含んだ状況型}.$$

$$(r21) \ll \text{英語約束型文}, [s \mid s \models \sigma] ; 1 \gg \Leftrightarrow \\ \ll \text{Promising}, [s \mid s \models \sigma] ; 1 \gg \mid \text{Promising に関する準備条件などを含んだ状況型}.$$

なお、上の表記において、記述を簡単にするために、事態の関係もパラメタとによってラベル付けした。また、次の記法を用いた。

$$(r22) \text{ReIn}(*R, * \sigma). \text{事態 } * \sigma \text{ が関係 } *R \text{ から構成される。}$$

(r23) Action(*R). 関係 *R が行為型である。

(r24) Appropriate-Value(*R, *role, *j). 対象 *j は、関係 *R の引数ロール *role の値として、適切である。

(r25) *hr | $\ll$ Hearer, *hr ;1 $\gg$. 現在の発話の聞き手 *hr.

(r26) *sp | $\ll$ Speaker, *sp ;1 $\gg$. 現在の発話の話し手 *sp.

(r27) Take-Part-in(*hr, * σ). 主体 *hr が、行為 * σ に参加可能である。

(r28) Agent(*hr, * σ). 主体 *hr が、行為 * σ の行為者である。

(r29) Unique(*R, *role, *x). 関係 *R の引数ロール *role として適切であり、他の対象とは区別できる対象 *x を含む状況の型。

制約 (r14), (r15) で用いられている状況の型は、第3.3章で述べた環境条件 (1) に対応している。制約 (r16), (r17) では、環境条件 (2) に対応する状況の型を利用している。制約 (r18), (r19) は、環境条件 (3) に対応している。また、制約 (r20), (r21) では、発話内行為の成立に関する準備条件を環境として利用しているが、これは、国際会議の問合せのような対話では、通常、成立しているものとして仮定する。

3.5 多様なレベルでの翻訳

翻訳の過程は、異なる自然言語の文を記述するための2つ表現が、ある環境に埋め込まれているとき、等価な情報内容をもつことを保証する過程である。すなわち、環境 C_1 に埋め込まれた原言語に属する表現 R_1 が与えられるとき、それを目標言語の表現に翻訳するとは、環境 C_2 に埋め込まれた同値な情報内容をもつ目標言語の表現 R_2 を見つけることである。この同値関係を次のように書く。

(r30) $R_1(*\vec{x}) \mid C_1(*\vec{x}) \cong R_2(*\vec{x}) \mid C_2(*\vec{x})$.

環境に関する条件は、必要のないときには省略する。例えば、日本語と英語の語彙のレベルの同値関係は次のように与えられる。

(r31) 銀行振り込み $\cong$ 'bank transfer'.

(r32) 支払う $\cong$ Pay.

上の同値関係 (r31), (r32) も実際には環境に依存することが考えられるが、ここでは無視する。また、同値関係 (r30) で言及されている環境条件 C_1, C_2 を利用することによって、効率的な表現を直接に介して、翻訳を行なうことができる。例えば、'支払う' と 'Pay' の対象格が明示されていない表現間の同値関係は、2つの表現が埋め込まれている各環境の中に、対象格となり得る翻訳的に同値なものが含まれているという条件を与えることによって保証される。この同値関係は、次のように書かれる。

(r33) $\ll$ 支払う, { AGEN $\rightarrow$ *x, MANNER $\rightarrow$ 銀行振り込み } ;1 $\gg$ | Unique(支払う, OBJE, *u) $\cong$
 $\ll$ Pay, { AGEN $\rightarrow$ *y, MANNER $\rightarrow$ 'bank transfer' } ;1 $\gg$ | Unique(Pay, OBJE, *v),
ただし、*x $\cong$ *y & *u $\cong$ *v.

例えば、対話 (d1), (d2) においては、第3.3章で述べた状況タイプ (r12) で分類される環境が導入されており、この環境は、同値関係 (r33) の環境条件を満たす。すなわち、翻訳プロセスの中で、対話 (d1), (d2) の文 (u4), (u7), (u12), (u15) を解釈、生成する際に '支払う (Pay)' の対象ロールの値を明示した表現 (r10), (r11) を介する必要はない。

また、次に示すように、この同値関係と制約を用いて、新たな同値関係を導くことができる。ただし、 $*\vec{x}$ は、パラメタのベクトルであり、 $\vec{A}$ は、定数項のベクトルである。ベクトルである。

(r34) IF $R_1(*\vec{x}) \Leftrightarrow T_1(*\vec{x}) \mid C_1(*\vec{x})$ &
 $R_2(*\vec{x}) \Leftrightarrow T_2(*\vec{x}) \mid C_2(*\vec{x})$ &
 $T_1(\vec{A}) \mid C_3(*\vec{x}) \cong T_2(\vec{A}) \mid C_4(*\vec{x})$ &
 $C_i(*\vec{x})$ が $*\vec{x} = \vec{A}$ の下で充足する ($i = 1 \cdots 4$)
THEN $R_1(\vec{A}) \cong R_2(\vec{A})$.

すなわち、原言語に属する表現 $R_1(\vec{A})$ を、それが適切な環境の下で用いられていることを保証することによって、表現 T_1, T_2 を介して目標言語の表現 $R_2(\vec{A})$ に翻訳することができる。このとき、翻訳に介在する表現として、様々な効率度のレベルの表現を選択することができ、必ずしも、文脈に依存する要因を排除したような完全な表現を得るまで解釈を進める必要はない。

例えば、対話 (d1) において、文 (u3) と (u6) との間の翻訳は、この発話の環境 (r12) の下に、制約 (r14), (r15) を用いて、表現 (r1) と、表現 (r9) とその英語側の表現、表現 (r2) とを介して、行うことができる。次に、文 (u3) と (u7) との間の翻訳について考える。このとき、第3.3で述べたように、'支払う (Pay)' 型の表現の行為者ロールの値は、待遇関係に関する環境から与えられるとして、ここでは無視する。環境 (r12) の下で、制

約 (r14), (r16), (r20) を用いて、表現 (r1) と (r9)、表現 (r7) とその英語側の表現、表現 (r5) とを介して、行なうことができる。この時、‘支払い’の対象を明示したような表現 (r10) は、この環境の下では必要ない。

4 関連研究

本枠組では、表現の情報内容を記述する際に、表現の情報内容をすべて表現自体に担わせるのではなく、表現が埋め込まれている環境に情報内容のある部分を担わせ、複雑な表現に頼ることなく、効率的な表現を直接扱うという方針を採っている。第2.3章で述べたように、この方針は、効率的な表現の文脈依存性に関して状況意味論で採られているアプローチ [1, Chapter 6 and 7] から大きな影響を受けている。実際、同様の立場に基づいて英語3人称代名詞の照応と限量子のスコープの分析が行なわれている [3]。本稿では、状況意味論における表現の情報内容と環境との間の関係についての見方が、省略文や断片文の解釈、生成、翻訳にも役立つことを示した。

また、本稿では、効率的な表現を直接扱い、それが適切に使用できる環境を分類するというアプローチを採った。従来の発話の解釈中心のモデルで利用されたプランやセンターという文脈情報は、もちろん発話の環境として利用できるが、本稿の重要な主張は、第2章で述べたように、省略文や断片文に対応する完全な文に直接対応付けることはしないで、個々の効率的な表現が適切に使用できるように環境に課せられる条件を記述する必要があるということである。単に発話者のもつプランが発話の環境であるというのは、断片文や省略文の使用を説明するためには漠然過ぎる条件である。

この枠組では、異なる効率度の表現を介して翻訳を行なうことができる。発想推論 (abductive reasoning) に基づいた翻訳モデル [4] が提案されているが、そのモデルも同様の利点をもっている。すなわち、そのモデルでは、適切な翻訳を行なうために必要なだけの発想推論を行ない、翻訳に不必要な情報を導き出さずに済ますことができる。確かに、発想推論という推論の枠組は、翻訳を行なうために必要なだけの推論を増進的に行なうモデルにとっては、適切なものである。しかし、その表現の枠組は、効率的な表現を直接に扱うという観点から構築されたものではない。本稿では、効率的な表現がある情報内容を運ぶために適切に使用できる環境を分類することによって、複雑で完全な表現に頼ることなしに、断片文や省略文の解釈 / 生成 / 翻訳を行なうことができることを示した。

5 おわりに

本稿においては、対話翻訳のプロセスの中で、いわゆる断片文や省略文と呼ばれる効率的な言語表現の解釈と生成のモデルを構築する際の基盤となる表現の枠組を示した。この枠組では、モデルの内部で効率的な表現を直接に取り扱うことができる。文の情報内容を特徴付けるために、効率的な言語表現とその環境を用意し、表現が適切な環境に埋め込まれているなら、効率度の異なる表現で同一の情報内容を記述できる。このことによって、不必要に深い解釈を行なって完全な表現を得る必要がなくなる。また、文脈の下に適切な断片文や省略文を生成することができる。さらに、この枠組は、環境に関する条件を参照しながら、多様な効率度のレベルの表現を介して、翻訳を行なうことを可能にする。

この枠組は、日英双方向の対話文翻訳に役立てることを狙いとしているが、本稿では、日英間対話翻訳の文脈の下に、主として、日本語断片文や省略文の解釈と生成に対して、枠組の適用を行なった。しかし、モデルの有効性を示すために援用した対話例は少数であるし、英語対話側の効率的な言語表現に関する考察も不十分である。今後、日本語断片文や省略文の使用に関するさらに包括的なモデルを本枠組に基づいて構築し、英語対話における効率的な言語表現、特に、代名詞を含んだ文と日本語対話文との間の翻訳プロセスに対してモデルを適用していく必要がある。さらに、代名詞を含む効率的な表現の解釈や生成には、統語的な制約も重要な役割を果たすが、本稿では、それについては言及しなかった。また、第3.1章の冒頭でも述べたように、対話の構造を抽出することによって、文を埋め込む環境を対話の流れの中から抽出する方法論を示す必要がある。これらは、今後の課題である。

参考文献

- [1] Barwise, J. *The Situation in Logic*, CSLI Lecture Notes 17 (1989).
- [2] 堂坂. 「対話登場人物を指示する日本語ゼロ代名詞の同定」, 「談話理解とその応用」シンポジウム, 情報処理学会 (1989).
- [3] Gawron, J. M. and S. Peters. *Anaphora and Quantification in Situation Semantics*, CSLI Lecture Notes 19 (1990).
- [4] Hobbs, J. R. and M. Kameyama. "Translation by abduction", COLING-90(1990), 155-161.
- [5] Kameyama, M. "A property-sharing constraint in centering", ACL-86(1986), 200-206.
- [6] 久野. 「談話の文法」, 大修館書店 (1978).
- [7] Litman, D. J. "Understanding plan ellipsis", AAAI-86(1986), 619-624.
- [8] Searle, J. *Expression and Meaning*, Cambridge University Press (1979).