Resolving Translation Mismatches With Information Flow

Megumi Kameyama, Ryo Ochitani*, Stanley Peters
The Center for the Study of Language and Information
Ventura Hall, Stanford University, Stanford, CA 94305

ABSTRACT

Languages differ in the concepts and real-world entities for which they have words and grammatical constructs. Therefore translation must sometimes be a matter of approximating the meaning of a source language text rather than finding an exact counterpart in the target language. We propose a translation framework based on Situation Theory. The basic ingredients are an information lattice, a representation scheme for utterances embedded in contexts, and a mismatch resolution scheme defined in terms of information flow. We motivate our approach with examples of translation between English and Japanese.

1 Introduction

The focus of machine translation (MT) technology has been on the translation of sentence structures out of context. This is doomed to limited quality and generality since the grammars of unlike languages often require different kinds of contextual information. Translation between English and Japanese is a dramatic one. The definiteness and number information required in English grammar is mostly lacking in Japanese, whereas the honorificity and speaker's perspectivity information required in Japanese grammar is mostly lacking in English. There are fundamental discrepancies in the extent and types of information that the grammars of these languages choose to encode.

An MT system needs to reason about the context of utterance. It should make adequate assumptions when the information required by the target language grammar is only implicit in the source language. It should recognize a particular discrepancy between the two grammars, and systematically react to the needs of the target language.

We propose a general reasoning-based model for handling translation mismatches. Implicit information is assumed only when required by the target language grammar, and only when the source language text allows it in the given context. Translation is thus viewed as a chain of reactive reasoning between the source and target languages.[1]

An MT system under this view needs: (a) a uniform representation of the context and content of utterances in discourse across languages, (b) a set of well-defined reasoning processes within and across languages based on the above uniform representation, and (c) a general treatment of translation mismatches.

In this paper, we propose a framework based on Situation Theory (Barwise and Perry 1983). First we will define the problem of translation mismatches, the key translation problem in our view. Second we will define the situated representation of an utterance. Third we will define our treatment of translation mismatches as a flow of information (Barwise and Etchemendy 1990). At the end, we will discuss a translation example.

2 What is a translation mismatch?

Consider a simple bilingual text:

```
    EXAMPLE 1:  BLOCKS (an AI problem)
ENGLISH:
Consider the blocks world with three blocks,
A, B, and C. The blocks A and B are on the table.
C is on A. Which blocks are clear?

JAPANESE:
3つの積木AとBとCがある積木の世界を考えてみる。
mittu no tumiki A to  B to  C ga   aru  tumiki no sekai wo  kangaete
three of block  A and B and C NOM exist block  of world ACC consider
miru
try
AとBの積木は机の上に乗っている。
A to  B no tumiki ha     tukue no ue    ni   notteiru
A and B of block  TOPIC table of above LOC riding
CはAの上に乗っている。
C ha     A no ue    ni   notteiru
C TOPIC A of above LOC riding
何も上に乗っていない積木はどれか。
nanimo ue    ni   notteinai tumiki ha     dore  ka
nothing above LOC riding    block  TOPIC which ?
```

Note the translation pair *C is on A* and CはAの上に乗っている (*C ha A no ue ni notteiru*). In En-

[1] Such a reasoning-based MT system is one kind of "negotiation"-based system, as proposed by Martin Kay. We thank him for stimulating our thinking.

*Linguistic Systems, Fujitsu Laboratories Ltd.

glish, the fact that C is on top of A is expressed using the preposition *on* and verb *is*. In Japanese, the noun 上 (*ue*) alone can mean either "on top of" or "above", and there is no word meaning just "on top of". Thus the Japanese translation narrows the relationship to the one that is needed by bringing in the verb 乗っている (*notteiru*) 'riding'. This phenomenon of the same information being attached to different morphological or syntactic forms in different languages is a well-recognized problem in translation.

TRANSLATION DIVERGENCES[2] of this kind manifest themselves at a particular representation level. They can be handled by (i) STRUCTURE-TO-STRUCTURE TRANSFERS, e.g., structural transformations of Nagao (1987), the sublanguage approach of Kosaka et al (1988), or by (ii) TRANSFER VIA A "DEEPER" COMMON GROUND, e.g., the entity-level of Carbonell and Tomita (1987), the lexical-conceptual structure of Dorr (1990). A solution of these types is not general enough to handle divergences at all levels, however. More general approaches to divergences allow (iii) MULTI-LEVEL MAPPINGS, i.e., direct transfer rules for mapping between different representation levels, e.g., structural correspondences of Kaplan et al. (1989), typed feature structure rewriting system of Zajac (1989), and abduction-based system of Hobbs and Kameyama (1990).

We want to call special attention to a less widely recognized problem, that of TRANSLATION MISMATCHES. They are found when the grammar of one language does not make a distinction required by the grammar of the other language. For instance, English noun phrases with COUNT type head nouns must specify information about definiteness and number (e.g. *a town, the town, towns*, and *the towns* are well-formed English noun phrases, but not *town*). Whereas in Japanese, neither definiteness nor number information is obligatory. Note the translation pair ***Which blocks are clear?*** and 何も上に乗っていない積木はどれか (***Nanimo ue ni notteinai tumiki ha dore ka***) above. ***Blocks*** is plural, but ***tumiki*** has no number information.

A mismatch has a predictable effect in each translation direction. From English into Japanese, the plurality information gets lost. From Japanese into English, on the other hand, the plurality information must be explicitly added.

Consider another example, a portion of step-by-step instructions for copying a file from a remote system to a local system:

EXAMPLE 2: FTP

```
ENGLISH:
2. Type 'open', a space, and the name of the
remote system, and press [return].
The system displays system connection messages
and prompts for a user name.
3. Type the user name for your account on the
remote system and press [return].
The system displays a message about passwords
and prompts for a password if one is required.
```

JAPANESE:

2. open 空白 リモートシステム名をタイプしリターンを押す。
'open' kuuhaku rimooto sisutemu mei wo taipu si [RETURN]
'open' space remote system name ACC type and [RETURN]
wo osu
ACC push

システム接続メッセージとユーザ名を問うプロンプトが表示される。
sisutemu setuzoku messeezi to yuusaa mei wo tou puronputo
system connection message and user name ACC ask prompt
ga hyouzi sareru
NOM display PASSIVE

3. リモートシステムでの自分のアカウントのユーザ名をタイプしリターンを押す。
rimooto sisutemu deno zibun no akaunto no yuuza mei
remote system LOC SELF of account of user name
wo taipu si [RETURN] wo osu
ACC type and [RETURN] ACC push

パスワードに関するメッセージと、もしパスワードが必要ならパスワードを問うプロンプトが表示される。
pasuwaado ni kansuru messeezi to, moshi pasuwaado ga
password about message and, if password NOM
hituyou nara passuwaado wo tou pronputo ga hyouji sareru
required then password ACC ask prompt NOM display PASSIVE

The notable mismatches here are the definiteness and number of the noun phrases for "space," "user name," "remote system," and "name" of the remote system in instruction step 2, and those for "message," "password," and "user name" in step 3. This information must be made explicit for each of these references in translating from Japanese into English whether or not it is decidable. It gets lost (at least on the surface) in the reverse direction.

Two important consequences for translation follow from the existence of major mismatches between languages. First in translating a source language sentence, mismatches can force one to draw upon information not expressed in the sentence—information only inferrable from its context at best. Secondly, mismatches may necessitate making information explicit which is only implicit in the source sentence or its context. For instance, the alternation of viewpoint between user and system in the FTP example is implicit in the English text, detectable only from the definiteness of noun phrases like "a/the user name" and "a password," but Japanese grammar requires an explicit choice of the user's viewpoint to use the reflexive pronoun *zibun*.

When we analyze what we called translation divergences above more closely, it becomes clear that divergences are instances of lexical mismatches. In the blocks example above, for instance, there is a mismatch between the spatial relations expressed with English *on*, which implies contact, and Japanese

[2] This term was taken from Dorr (1990) where the problem of divergences in verb predicate-argument structures was treated. Our use of the term extends the notion to cover a much more general phenomenon.

ue, which implies nothing about contact. It so happens that the verb "notteiru" can naturally resolve the mismatch within the sentence by adding the information "on top of". Divergences are thus lexical mismatches resolved within a sentence by coocurring lexemes. This is probably the preferred method of mismatch resolution, but it is not always possible. The mismatch problem is more dramatic when the linguistic resources of the target language offer no natural way to match up with the information content expressed in the source language, as in the above example of definiteness and number. This problem has not received adequate attention to our knowledge, and no general solutions have been proposed in the literature.

Translation mismatches are thus a key translation problem that any MT system must face. What are the requirements for an MT system from this perspective? First, mismatches must be made recognizable. Second, the system must allow relevant information from the discourse context be drawn upon as needed. Third, it must allow implicit facts be made explicit as needed. Are there any systematic ways to resolve mismatches at all levels? What are the relevant parameters in the "context"? How can we control contextual parameters in the translation process? Two crucial factors in an MT system are then REPRESENTATION and REASONING. We will first describe our representation.

3 Representing the translation content and context

Translation should preserve the information content of the source text. This information has at least three major sources: Content, Context, Language. From the content, we obtain a piece of information about the relevant world. From the context, we obtain discourse-specific and utterance-specific information such as information about the speaker, the addressee, and what is salient for them. From the linguistic forms (i.e., the particular words and structures), through shared cooperative strategies as well as linguistic conventions, we get information about how the speaker intends the utterance to be interpreted.

DISTRIBUTIVE LATTICE OF INFONS. In this approach, pieces of information, whether they come from linguistic or non-linguistic sources, are represented as ***infons*** (Devlin 1990). For an n-place relation P, $\langle\langle P, x_1, ..., x_n ;1\rangle\rangle$ denotes the informational item, or infon, that $x_1, ..., x_n$ stand in the relation P, and $\langle\langle P, x_1, ..., x_n ;0\rangle\rangle$ denotes the infon that they do not stand in the relation. Given a situation s, and an infon σ, $s \models \sigma$ indicates that the infon σ is made factual by the situation s, read s *supports* σ .

Infons are assumed to form a distributive lattice with least element **0**, greatest element **1**, set I of infons, and "involves" relation $\Rightarrow$ satisfying:[3]

for infons σ and τ, if $s \models \sigma$ and $\sigma \Rightarrow \tau$ then $s \models \tau$

This distributive lattice $\langle I, \Rightarrow\rangle$, together with a nonempty set *Sit* of ***situations*** and a relation $\models$ on *Sit* $\times$ I constitute an ***infon algebra*** (see Barwise and Etchemendy 1990).

THE LINGUISTIC INFON LATTICE. We propose to use infons to uniformly represent information that come from multiple "levels" of linguistic abstraction, e.g., morphology, syntax, semantics, and pragmatics. Linguistic knowledge as a whole then forms a distributive lattice of infons.

For instance, the English words ***painting***, ***drawing***, and ***picture*** are associated with properties; call them P1, P2, and P3, respectively. In the following sublattice, a string in English (EN) or Japanese(JA) is linked to a property with the SIGNIFIES relation (written ==),[4] and properties themselves are interlinked with the INVOLVES relation ($\Rightarrow$):

EN: "picture" == P1⟨⟨picture, x; 1⟩⟩
EN: "painting" == P2⟨⟨painting, x; 1⟩⟩
EN: "drawing" == P3⟨⟨drawing, x; 1⟩⟩
EN: "oil painting" == P4⟨⟨oil painting, x; 1⟩⟩
EN: "water-color" == P5⟨⟨water-color, x; 1⟩⟩
P2 $\Rightarrow$ P1, P3 $\Rightarrow$ P1, P4 $\Rightarrow$ P2, P5 $\Rightarrow$ P2

So far the use of lattice appears no different from familiar semantic networks. Two additional factors bring us to the basis for a general translation framework. One is multi-linguality. The knowledge of any new language can be added to the given lattice by inserting new infons in appropriate places and adding more instances of the "signifies" relations. The other factor is grammatical and discourse-functional notions. Infons can be formed from any theoretical notions whether universal or language-specific, and placed in the same lattice.

Let us illustrate how the above "picture" sublattice for English would be extended to cover Japanese words for pictures. In Japanese, 絵 (*e*) includes both paintings and drawings, but not photographs. It is thus more specific than ***picture*** but more general than ***painting*** or ***drawing***. No Japanese words cosignify with ***painting*** or ***drawing***, but more specific concepts have words— 油絵 (***aburae***) for P4, 水彩画 (***suisaiga***) for P5, and the rarely used word 線描 (***senbyou***) for artists' line drawings. Note that synonyms co- signify the same property. (See Figure 1 for the extended sublattice.)

[3] We assume that the relation $\Rightarrow$ on infons is transitive, reflexive, and anti-symmetric after Barwise and Etchemendy.

[4] This is our addition to the infon lattice. The SIGNIFIES relation links the SIGNIFIER and SIGNIFIED to form a SIGN (de Saussure 1959). Our notation abbreviates standard infons, e.g., ⟨⟨signifies, "picture", EN, P1; 1⟩⟩ .

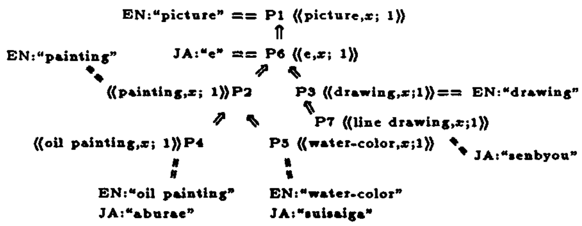

Figure 1: The "Picture" Sublattice

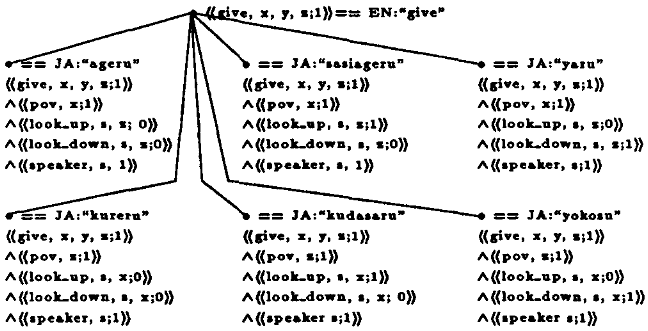

Figure 2: Verbs of giving

JA: "絵(*e*)" == P6((e, x; 1))
JA: "油絵(*aburae*)" == P4((oil painting, x; 1))
JA: "水彩画(*suisaiga*)" == P5((water-color, x; 1))
JA: "線描(*senbyou*)" == P7((senbyou, x; 1))
P2 ⇒ P6, P3 ⇒ P6, P6 ⇒ P1, P7 ⇒ P3

Lexical differences often involve more complex pragmatic notions. For instance, corresponding to the English verb ***give***, Japanese has six basic verbs of giving, whose distinctions hinge on the speaker's perspectivity and honorificity. For "X gave Y to Z" with neutral honorificity, ***ageru*** has the viewpoint on X, and ***kureru***, the viewpoint on Z. ***Sasiageru*** honors Z with the viewpoint on X, and ***kudasaru*** honors X with the viewpoint on Z, and so on. See Figure 2.

As an example of grammatical notions in the lattice, take the syntactic features of noun phrases. English distinguishes six types according to the parameters of count/mass, number, and definiteness, whereas Japanese noun phrases make no such syntactic distinctions. See Figure 3. Grammatical notions often draw on complex contextual properties such as "definiteness", whose precise definition is a research problem on its own.

THE SITUATED UTTERANCE REPRESENTATION. A translation should preserve as far as practical the information carried by the source text or discourse. Each utterance to be translated gives information about a situation being described—precisely what information depends on the context in which the utterance is embedded. We will utilize what we call a SITUATED UTTERANCE REPRESENTATION (SUR) to integrate the form, content, and

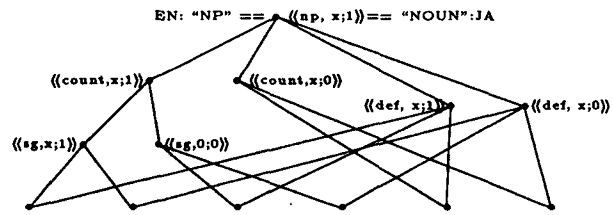

Figure 3: The "NP" Sublattice

context of an utterance.[5] In translating, contextual information plays two key roles. One is to reduce the number of possible translations into the target language. The other is to support reasoning to deal with translation mismatches.

Four situation types combine to define what an utterance is:

Described Situation The way a certain piece of reality is, according to the utterance

Phrasal Situation The surface form of the utterance

Discourse Situation The current state of the on-going discourse when the utterance is produced

Utterance Situation The specific situation where the utterance is produced

The content of each utterance in a discourse like the Blocks and FTP examples is that some situation is described as being of a certain type. This is the information that the utterance carries about the DESCRIBED SITUATION.

The PHRASAL SITUATION represents the *surface form* of an utterance. The orthographic or phonetic, phonological, morphological, and syntactic aspects of an utterance are characterized here.

The DISCOURSE SITUATION is expanded here in situation theory to characterize the dynamic aspect of discourse progression drawing on theories in computational discourse analysis. It captures the linguistically significant parameters in the current state of the on-going discourse,[6] and is especially useful for finding functionally equivalent referring expressions between the source and target languages.[7]

- reference time = the time pivot of the linguistic

[5] Our characterization of the context of utterance draws on a number of existing approaches to discourse representation and discourse processing, most notably those of Grosz and Sidner (1986), Discourse Representation Theory (Kamp 1981, Heim 1982), Situation Semantics (Barwise and Perry 1983, Gawron and Peters 1990), and Linguistic Discourse Model (Scha and Polanyi 1988).

[6] Lewis (1979) discussed a number of such parameters in a logical framework.

[7] Different forms of referring expressions (e.g. pronouns, demonstratives) and surface structures (i.e. syntactic and

description ("then")[8]

- point of view = the individual from whose viewpoint a situation is described[9]
- attentional state = the entities currently in the focus and center of attention[10]
- discourse structural context = where the utterance is in the structure of the current discourse[11]

The specific UTTERANCE SITUATION contains information about those parameters whose values support indexical references and deixes: e.g., information about the speaker, hearer(s), the time and location of the utterance, the perceptually salient context, etc.

The FTP example text above describes a situation in which a person is typing commands to a computer and it is displaying various things. Specifically, it describes the initial steps in copying a file from a remote system to a local system with ftp. Consider the first utterance in instruction step 3 repeated here: `Type the user name for your account on the remote system and press [return]`. It occurs in a type of DISCOURSE SITUATION where there has previously been mention of a remote system and where a pattern has been established of alternating the point of view between the addressee and another agent (the local computer system). We enumerate below some of the information in the SUR associated with this utterance.

The Described Situation (DeS) of the utterance is

≪type, y, n, t'; 1 ≫ ∧ ≪press, y, k, t''; 1 ≫ where n satisfies $n = n'$ ⟺ ≪named, a, n'; 1 ≫ a satisfies ≪account, a, y, r; 1 ≫ r satisfies ≪system, r; 1 ≫ ∧ ≪remote from, r, y; 1 ≫ t' satisfies ≪later, t', t; 1 ≫ k satisfies ≪named,k,[return];1≫ t'' satisfies ≪later, t'', t'; 1 ≫

The Phrasal Situation (PS) of the utterance is

≪language, u,English; 1 ≫ ∧ ≪written, u, "Type the user name for your account on the remote system and press [return]."; 1 ≫ ∧ ≪syntax, u,{...≪written, e, "the user name"; 1 ≫ ∧ ≪np, e; 1 ≫ ∧ ≪definite, e; 1 ≫ ∧ ≪singular, e; 1 ≫ ∧ ...}; 1 ≫

The Discourse Situation (DiS) is

$r = r'$ ⟺ ≪focus, r',remote system; 1 ≫

Finally, the Utterance Situation (US) is

≪uttering, x, u, t; 1 ≫ ∧ ≪addressing, x, y, t; 1 ≫

Note that the parameter y of DeS for the user (to whom the discourse is addressed) has its value constrained in US; the same is true of the parameter t for utterance time. Similarly, the parameter r of DeS for the definite remote system under discussion is assigned a definite value only by virtue of the information in DiS that it is the unique remote system that is salient at this point in the discourse. This cross-referencing of parameters between types constitutes further support for combining all four situation types in a unified SUR. In order for the analysis and generation of an utterance to be associated with an SUR, the grammar of a language should be a set of constraints on mappings among the values assigned to these parameters.

4 Translation as information flow

Translation must often be a matter of approximating the meaning of a source language text rather than finding an exact counterpart in the target language since languages differ in the concepts and real-world entities for which they have words and grammatical constructs.

In the cases where no translation with exactly the same meaning exists, translators seek a target language text that accurately describes the same real world situations as the source language text.[12] The situation described by a text normally includes additional facts besides those the text explicitly states. Human readers or listeners recognize these additional facts by knowing about constraints that hold in the real world, and by getting collateral information about a situation from the context in which a description is given of it. For a translation to be a good approximation to a source text, its "fleshed out" set of facts—the facts its sentences explicitly state plus the additional facts that these entail by known real-world constraints—should be a maximal subset of the "fleshed out" source text facts.

Finding a translation with the desired property can be simplified by considering not sets of facts (infons) but infon lattices ordered by involvement relations including known real-world constraints. If a given infon is a fact holding in some situation, all infons in such a lattice higher than the given one (i.e., all further infons it involves) must also be facts in the situation. Thus a good translation can be found by looking for the lowest infons in the lattice that the source text either explicitly or implicitly requires to hold in the described situation, and finding a target language text that either explicitly or implicitly requires the maximal number

---

phonetic) often carry equivalent discourse functions, so explicit discourse representation is needed in translating these forms. See also Tsujii (1988) for this point.

[8] Reichenbach (1947) pointed out the significance of reference time, which in the FTP example accounts for why the addressee is to press [return] after typing the user name of his/her remote account.

[9] Katagiri (to appear) describes how this parameter interacts with Japanese grammar to constrain use of the reflexive pronoun *zibun*.

[10] See Grosz (1977), Grosz et al. (1983), Kameyama (1986), Brennan et al. (1987) for discussions of this parameter.

[11] This parameter may be tied to the "intentional" aspect of discourse as proposed by Grosz and Sidner (1986). See, e.g., Scha and Polanyi (1988) and Hobbs (1990) for discourse structure models.

[12] In some special cases, translation requires mapping between different but equivalent real world situations, e.g., cars drive on different sides of the street in Japan and in the US.

of them to hold.[13]

THE INFORMATION FLOW GRAPH. Translation can be viewed as a flow of information that results from the interaction between the grammatical constraints of the source language (SL) and those of the target language (TL). This process can be best modelled with information flow graphs (IFG) defined in Barwise and Etchemendy 1990. An IFG is a semantic formalization of valid reasoning, and is applicable to information that comes from a variety of sources, not only linguistic but also visual and other sensory input (see Barwise and Etchemendy 1990b). By modelling a treatment of translation mismatches with IFGs, we aim at a semantically correct definition that is open to various implementations.

IFGs represent five basic principles of information flow:

Given Information present in the initial assumptions, i.e., an initial "open case."

Assume Given some open case, assume something extra, creating an open subcase of the given case.

Subsume Disregard some open case if it is subsumed by other open cases, any situation that supports the infons of the subsumed case supports those of one of the subsuming cases.

Merge Take the information common to a number of open cases, and call it a new open case.

Recognize as Possible Given some open case, recognize it as representing a genuine possibility, provided the information present holds in some situation.

RESOLVING MISMATCHES. First, a translation mismatch is recognized when the generation of a TL string is impossible from a given set of infons. More specifically,

> given a Situated Utterance Representation (SUR), when no phrasal situations of TL support SUR because no string of TL signifies infon σ in SUR, The TL grammar cannot generate a string from SUR, and there is a TRANSLATION MISMATCH on σ.

A translation mismatch on σ above is resolved in one of two directions:

> **Mismatch Resolution by Specification:** Assume a specific case τ such that $\tau \Rightarrow \sigma$ and there is a Phrasal Situation of TL that supports τ. A new open case SUR' is then generated, adding τ to SUR.

This is the case when the Japanese word 絵 (e) is translated into either *painting* or *drawing* in English. The choice is constrained by what is known in the given context.

> **Mismatch Resolution by Generalization:** Assume a general case τ such that $\sigma \Rightarrow \tau$ and there is a Phrasal Situation of TL that supports τ. A new open case SUR' is then generated, adding τ to SUR.

This is the case when the Japanese word 絵 (e) is translated into *picture* in English, or English words p*painting* and *drawing* are both translated into 絵 (e) in Japanese. That is, two different utterances in English, *I like this painting* and *I like this drawing*, would both be translated into 私はこの絵が好きです (*watasi wa kono e ga suki desu*) in Japanese according to this scheme.

Resolution by generalization is ordinarily less constrained than resolution by specification, even though it can lose information. It should be blocked, however, when generalizing erases a key contrast from the content. For example, given an English utterance, *I like Matisse's drawings better than paintings*, the translation into Japanese should not generalize both *drawings* and *paintings* into 絵 (e) since that would lose the point of this utterance completely. The mismatches must be resolved by specification in this case, resulting in, for instance, 私はマティスの油絵や水彩画よりも線描が好きです (*watasi wa Matisse no aburae ya suisaiga yorimo senbyou ga suki desu*) 'I like Matisse's line_drawings(P7) better than oil_paintings(P4) or water-colors(P5)'.

There are IFGs for the two types of mismatch resolution. Using ∘ for an open (unsubsumed) node and • for a subsumed node, we have the following:

Mismatch Resolution by Specification: (given $\tau \Rightarrow \sigma$)

Given: $\circ\{\sigma\}$ Assume: $\circ\{\sigma\}$ — $\circ\{\sigma, \tau\}$

Mismatch Resolution by Generalization: (given $\sigma \Rightarrow \tau$)

Given: $\circ\{\sigma\}$ Assume: $\circ\{\sigma\}$ — $\circ\{\sigma, \tau\}$ Subsume: $\bullet\{\sigma\}$ — $\circ\{\sigma, \tau\}$

Both resolution methods add more infons to the given SUR by ASSUMPTION, but there is a difference. In resolution by specification, subsequent subsumption does not always follow. That is, only by contradicting other given facts, can some or all of the newly assumed SUR's later be subsumed, and only by exhaustively generating all its subcases, the original SUR can be subsumed. In resolution by generalization, however, the newly assumed general case immediately subsumes the original SUR.[14]

[13] As more sophisticated translation is required, we could make use of the multiple situation types to give more importance to some aspects of translation than others depending on the purpose of the text (see Hauenschild (1988) for such translaion needs).

[14] Resolution by specification models a form of abductive inference, and generalization, a form of deductive inference

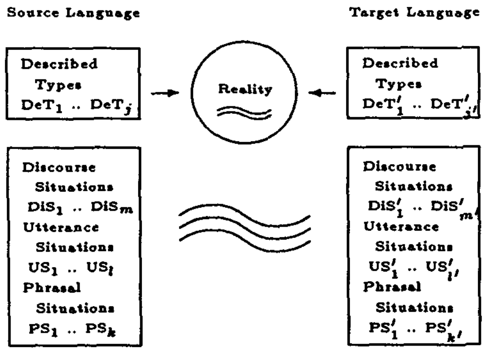

Figure 4: Situated Translation

THE TRANSLATION MODEL. Here is our characterization of a TRANSLATION:

> Given a SUR ⟨ DeT, PS, DiS, US ⟩ of the nth source text sentence and a discourse situation DiS″ characterizing the target language text following translation of the $(n-1)$st source sentence, find a SUR ⟨ DeT′, PS′, DiS′, US′⟩ allowed by the target language grammar such that DiS″ $\subseteq$ DiS′ and
>
> ⟨ DeT, PS, DiS, US ⟩ $\sim$ ⟨ DeT′, PS′, DiS′, US′⟩.
>
> ($\sim$ is the approximates relation we have discussed, which constrains the flow of information in translation.)

Our approach to translation combines SURs and IFGs (see Figure 4). Each SUR for a possible interpretation of the source utterance undergoes a FLOW OF TRANSLATION as follows: A set of infons is initially GIVEN in an SUR. It then grows by mismatch resolution processes that occur at multiple sites until a generation of a TL string is RECOGNIZED AS POSSIBLE. Each mismatch resolution involves ASSUMING new SUR's and SUBSUMING inconsistent or superfluous SUR's.[15]

Our focus here is the epistemological aspect of translation, but there is a heuristically desirable property as well. It is that the proposed mismatch resolution method uses only so much additional information as required to fill the particular distance between the given pair of linguistic systems. That is, the more similar two languages, less computation. This basic model should be combined with various control strategies such as default reasoning in an actual implementation.

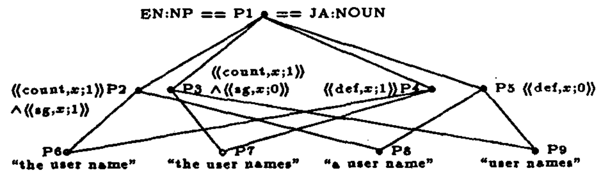

Figure 5: The IFG for NP Translation

5 A translation example

We will now illustrate the proposed approach with a Japanese-to-English translation example: the first sentence of instruction step 3 in the FTP text.

INPUT STRING: "3. リモートシステムでの自分のアカウントのユーザ名をタイプしリターンを押す。"

1. In the initial SUR are infons for リモートシステム (*rimooto sisutemu*) 'remote system', アカウント (*akaunto*) 'account', and ユーザ名 (*yuuzaa mei*) 'user name'. All of thesewords signify properties that are signified by English COUNT nouns but the Japanese SUR lacks definiteness and number information.
2. Generation of English from the SUR fails because, among other things, English grammar requires NPs with COUNT head nouns to be of the type, Sg-Def, Sg-Indef, Pl-Def, or Pl-Indef. (translation mismatch)
3. This mismatch cannot be resolved by generalization. It is resolved by assuming four subcases for each nominal, and subsuming those that are inconsistent with other given information. The "remote system" is a singular entity in focus, so it is Sg-Def, and the other three subcases are subsumed. The "user name" is an entity in center, so Definite. The "account" is Definite despite its first mention because its possesser (addressee) is definite. Both "user name" and "account" can be either Singular or Plural at this point. Let's assume that a form of default reasoning comes into play here and concludes that a user has only one user name and one account name in each computer.
4. The remaining open case permits generation of English noun phrases, so the translation of this utterance is done.

OUTPUT STRING: "Type the user name for your account on the remote system and ..."

6 Conclusions

In order to achieve high-quality translation, we need a system that can reason about the context of utterances to solve the general problem of transla-

in a situation-theoretic context. One way to implement these methods is in the abduction-based system proposed by Hobbs and Kameyama (1990).

[15] A possible use of MERGE in this application is that two different SUR's may be merged when an identical TL string would be generated from them.

tion mismatches. We have proposed a translation framework based on Situation Theory that has this desired property. The situated utterance representation of the source string embodies the contextual information required for adequate mismatch resolution. The translation process has been modelled as a flow of information that responds to the needs of the target language grammar. Reasoning across and beyond the linguistic levels, this approach to translation respects and adapts to differences between languages.

7 Future implications

We plan to design our future implementation of an MT system in light of this work. Computational studies of distributive lattices constrained by multiple situation types are needed. Especially useful linguistic work would be on grammaticized contextual information. More studies of the nature of translation mismatches are also extremely desirable.

The basic approach to translation proposed here can be combined with a variety of natural language processing frameworks, e.g., constraint logic, abduction, and connectionism. Translation systems for multi-modal communication and those of multiple languages are among natural extensions of the present approach.

8 Acknowledgements

We would like to express our special thanks to Hidetoshi Sirai. Without his enthusiasm and encouragement at the initial stage of writing, this paper would not even have existed. This work has evolved through helpful discussions with a lot of people, most notably, Jerry Hobbs, Yasuyoshi Inagaki, Michio Isoda, Martin Kay, Hideo Miyoshi, Hiroshi Nakagawa, Hideyuki Nakashima, Livia Polanyi, and Yoshihiro Ueda. We also thank John Etchemendy, David Israel, Ray Perrault, and anonymous reviewers for useful comments on an earlier version.

References

[1] Barwise, Jon and John Etchemendy. 1990. Information, Infons, and Inference. In Cooper et al. (eds), 33-78.

[2] Barwise, Jon and John Etchemendy. 1990b. Visual Information and Valid Reasoning. In W. Zimmerman (ed.) *Visualization in Mathematics*. Washington DC: Mathematical Association of America.

[3] Barwise, Jon, and John Perry. 1983. *Situations and Attitudes*. Cambridge, MA: MIT Press.

[4] Brennan, Susan, Lyn Friedman, and Carl Pollard. 1987. A Centering Approach to Pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA: ACL, 155-162.

[5] Carbonell, Jaime G. and Masaru Tomita. 1987. Knowledge-based Machine Translation, the CMU Approach. In Nirenburg (ed.), 68-89.

[6] Cooper, Robin, Kuniaki Mukai, and John Perry (eds) 1990. *Situation Theory and Its Applications*, Volume 1, CSLI Lecture Notes Number 22. Stanford: CSLI Publications.

[7] Devlin, Keith. 1990. Infons and Types in an Information-Based Logic. In Cooper et al. (eds), 79-96.

[8] Dorr, Bonnie. 1990. Solving Thematic Divergences in Machine Translation. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, Pittsburgh, PA, 127-134.

[9] Gawron, J. Mark and Stanley Peters. 1990. *Anaphora and Quantification in Situation Semantics*, CSLI Lecture Notes Number 19. Stanford: CSLI Publications.

[10] Grosz, Barbara. 1977. The Representation and Use of Focus in Dialogue Understanding. Technical Report 151, SRI International, Menlo Park, CA.

[11] Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1983. Providing a Unified Account of Definite Noun Phrases in Discourse. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA, 44-50.

[12] Grosz, Barbara J. and Candace L. Sidner. 1986. Attention, Intention, and the Structure of Discourse. *Computational Linguistics*, 12(3), 175-204.

[13] Hauenschild, Christa. 1988. Discourse Structure – Some Implications for Machine Translation. In Maxwell et al. (eds), 145-156.

[14] Heim, Irene R. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD dissertation, University of Massachusetts at Amherst.

[15] Hobbs, Jerry. 1990. *Literature and Cognition*. CSLI Lecture Note Number 21. Stanford: CSLI Publications.

[16] Hobbs, Jerry and Megumi Kameyama. 1990. Translation by Abduction. In *Proceedings of the 13th International Conference on Computational Linguistics*, Helsinki, Finland.

[17] Kameyama, Megumi. 1986. A Property-sharing Constraints in Centering. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA: ACL, 200-206.

[18] Kamp, Hans. 1981. A Theory of Truth and Semantic Representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds), *Formal Methods in the Study of Language*. Amsterdam: Mathematical Center.

[19] Kaplan, Ronald M., Klaus Netter, Jürgen Wedekind, and Annie Zaenen. 1989. Translation by Structural Correspondences. In *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics*, Manchester, United Kingdom, 272-281.

[20] Katagiri, Yasuhiro. To appear. Structure of Perspectivity: A Case of Japanese Reflexive Pronoun "zibun". Paper presented at STASS-90, Scotland.

[21] Kosaka, Michiko, Virginia Teller, and Ralph Grishman. 1988. A Sublanguage Approach to Japanese-English Machine Translation. In Maxwell et al. (eds), 109-122.

[22] Lewis, David K. 1979. Scorekeeping in Language Game. In Bäuerle, R., U. Egli and A. von Stechow (eds) *Semantics from Different Points of View*. Berlin: Springer Verlag.

[23] Maxwell, Dan, Klaus Schubert, and Toon Witkam (eds). 1988. *New Directions in Machine Translation*. Dordrecht, Holland: Foris.

[24] Nagao, Makoto. 1987. The Role of Structural Transformation in a Machine Translation System. In Nirenburg (ed.), 262-277.

[25] Nirenburg, Sergei (ed.) 1987. *Machine Translation*. Cambridge: Cambridge University Press.

[26] Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York: Dover.

[27] de Saussure, Ferdinand. 1959. *Course in General Linguistics*. Edited by Charles Bally and Albert Sechehaye in collaboration with Albert Riedlinger. Translated by Wade Baskin. New York: McGraw-Hill.

[28] Scha, Remko and Livia Polanyi. 1988. An Augmented Context-free Grammar of Discourse. In *Proceedings of the 12th International Conference on Computational Linguistics*, Budapest, Hungary.

[29] Tsujii, Jun-ichi. 1988. What is a Cross-linguistically Valid Interpretation of Discourse? In Maxwell et al. (eds), 157-166.

[30] Zajac, Remi. 1989. A Transfer Model Using a Typed Feature Structure Rewriting System with Inheritance. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, Vancouver, Canada.