Meaning preservation in Machine Translation

Michael Carl
Institut für Angewandte Informationsforschung,
Martin-Luther-Straße 14,
66111 Saarbrücken, Germany,
carl@iai.uni-sb.de

Abstract

The paper investigates how meaning preservation can be achieved in MT-systems. A distinction is made between *ontological* MT-systems and *epistemological* MT-systems. Whereas ontological MT-systems perform meaning preservation in the translation process by means of a set of rules that is provided by the system-external world, epistemological MT-systems perform an 'understanding' of meaning that is induced through a learning corpus. Theories of meaning, which are implemented by MT-systems, can be *rich* or they can be *austere* and they can be *holistic* or *molecular*.
Some approaches to epistemological MT are discussed and classified according to the terminology introduced. The paper states that "all-purpose" MT is, however, unfeasible.

In the MT literature it has often been stated that a translation is valid if and only if the source language text and the target language text have the same meaning. Although this appears intuitively correct, it is unclear what meaning actually is: many people consider meaning as a mental phenomenon. For a great deal of research in the field of computational linguistics and artificial intelligence meaning is merely a (bracketed) expression that is part of a (meta) language. Some philosophers conceptualize meaning as a compound structure from which parts are in their minds and other parts are in a (physical) world[1]. Others see it as an abstract entity that is shared by the members of a (linguistic) society[2] and still others define the meaning as being dependent on the goals and intentions of the cognitive agent: meaning thus becomes a concept which is relative to an agent's theory (about the world) [Sch94].
I do not want to discuss the question of whether the meaning (or parts of it) are in the mind or in the world, rather I will investigate how the meaning preservation requirement is handled in MT-systems. I will use the following notations:

- If a source text *source* and a target text *target* are translations of each other I write $T_{source} \leftrightarrow T_{target}$.
- If a source text *source* and a target text *target* have an equivalent meaning I write $M_{source} \equiv M_{target}$.

we thus obtain: $T_{source} \leftrightarrow T_{target}$ is true iff $M_{source} \equiv M_{target}$ or, equally[3]:

$$T_{source} \leftrightarrow T_{target} \Longleftrightarrow M_{source} \equiv M_{target} \qquad (1)$$

The above definition is similar to the Tarski definition of truth [Tar35]: "T" is true if and only if T, where T can be replaced by any sentence. Thus, "snow is white" is true if and only if snow is white.
This approach known as 'disquotation' method to truth semantics can equally be applied to translation: $T_{\text{snow is white}} \leftrightarrow T_{\text{Schnee ist weiss}}$ is true if and only if $M_{\text{snow is white}} \equiv M_{\text{Schnee ist weiss}}$. Note that there is still a certain degree of uncertainty in the above equivalence. Thus, $T_{source} \leftrightarrow T_{target}$ can be true (T_{target} can be a translation for T_{source}) even if the meanings of both texts are unknown. However, I will exclude this possibility as a feasible way for MT since, if the meaning of a text is unknown, the translation does not appear to be computable.
MT-systems have to tackle with the meaning preservation of the source and the (generated) target text (otherwise they would not be MT-systems). However, very different approaches can be found.
In this paper I will distinguish *ontological* MT approaches from *epistemological* MT approaches. If the former MT-systems implement a theory of meaning which is formulated through a system-external component that belongs to the WORLD[4], the latter approaches implement a program to understand the meaning preservation requirements of the system-external WORLD.

Ontological Machine Translation

Ontological MT-systems (OMT) have a fixed set of rules that map an input text onto a meaning preserving output text. The validity of the transformations is checked by an Evaluation process outside the system (i.e. in the WORLD) which — in case of necessity — modifies the mapping rules to enhance the meaning preservation capacities of the system. In order for the Evaluation process to formulate the appropriate mapping rules, it must have

[1] Putnam [Put75] defines meaning as a four dimensional vector containing the 1. syntactic marker 2. semantic marker 3. stereotypes, and 4. the extension of the word.

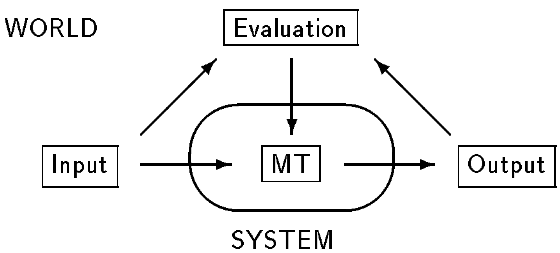
[2] Frege [Fre62] differentiates imagination (Vorstellung), meaning (Sinn) and reference (Bedeutung). Whereas the imagination of a word or sentence is subjective, its meaning is a social phenomenon and the reference is in the world.

[3] Note that this definition does not assume any theory of meaning. In particular it leaves open in what consists an 'equivalence of meaning'

[4] In this paper I use in capital letters WORLD to refer to everything outside the MT-system.

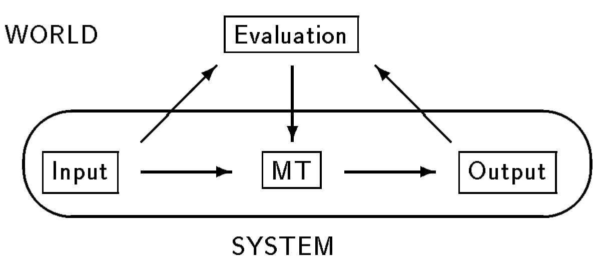
a theory of meaning, which determines the required mapping rules. In such an Ontological Machine Translation scenario (OMT-1 according to Figure-1 below), it is claimed that the Evaluation process can find in a finite elapse of time a set of rules that map an arbitrary input text onto a (meaning preserving) output text.

Figure 1
OMT-1: The MT-system in the oval frame translates arbitrary input text into a meaning preserving output text. The Evaluation *process in the* WORLD *provides the necessary set of mapping rules*



From the system's point of view, with each new set of rules, it acquires more (or — at least — different) knowledge about how to map an input source text onto an output target text. The system thus acquires ontological knowledge about an outside WORLD.
In order to enable the Evaluation to formulate such a set of rules, much effort has been given to find an ontological theory of meaning. Such concepts as "deep semantic structure", "(linguistic) universals" and the like are part of this research. However, for unrestricted texts and arbitrary languages none of these concepts are generally accepted.

Figure 2
OMT-2: The MT-system in the oval frame translates restricted input text for which it was designed. The system external Evaluation *provides the necessary set of mapping rules*



Some MT-approaches, therefore, restrict the input text to a certain domain. The METEO 96 [CG96] MT-system, for instance, is only designed to translate weather-reports from English to French. Instead of 'knowing' the meaning of the whole linguistic universe, in such an Ontological Machine Translation scenario (OMT-2) as illustrated in Figure 2, only a restricted type of input texts is accepted for which the system was designed.
In the same way there is no single type of vehicle that serves all purposes such as family excursions, racing or construction material transportation, rarely someone expect to have a single MT-system that serves such different purposes as scientific translation, instruction manual translation or even newspaper translation.
In an ontological MT-system, the meaning preserving mapping rules are given by an outside component in the WORLD. Meaning preservation in the translation process is — for the system — a matter of executing a set mapping rules which are expressions of a theory of meaning formulated by the Evaluation process.

Why OMT-1 cannot work

In the OMT-1 scenario, it is claimed that a finite set of rules can map an arbitrary input text onto a meaning preserving output text. This would only be possible if the Evaluation process can formulate an appropriate (i.e. *full blooded*) theory that leads to an understanding of all types of input text and that can be communicated to the system.
If we had such a full blooded theory of meaning, obviously we could implement an all-purpose MT-system because the theory would give us the necessary means to do so. On the other hand, if we had an all-purpose MT-system, we could generate such a theory simply by taking the output of the MT-system as expressions of that theory. In this latter case, for each expression of the language, the system would generate the appropriate expression of the metalanguage. All-purpose MT-systems and full blooded theories of meaning are thus equivalent.
In this section I shall show that it is impossible to formulate a full blooded theory of meaning and that at best OMT-2 can be realized.
Dummett [Dum75] differentiates three types of theories of meaning:

- A *full blooded* theory of meaning leads to an understanding of the object language without making use of another language that requires the concepts to be already known. It thus fully explains all concepts in a language and their constituting primitives. A full blooded theory of meaning explains " 'X' means X' where the metalanguage does not assume any understanding of the primitives contained in X.
- A *modest* theory of meaning leads to an understanding of the object language via a grasp of the concepts expressed by its primitive expressions. It presupposes, thus, an understanding of these primitives. A modest theory of meaning explains " 'X' means X' where the metalanguage assumes an understanding of the primitives contained in X.
- A *translation manual* leads to an understanding of the translated language via an understanding of the language into which the translation is made. A translation manual merely states that '$X \longleftrightarrow Y$' where the understanding of X results from an understanding of Y.

In contrast to a full blooded theory of meaning, neither a translation manual nor a modest theory of meaning fully displays what an understanding of the object language consists in. A translation manual presupposes the mastery of the target language to derive the understanding of the translated language. A modest theory presupposes

the knowledge of the propositions that are expressed in the metalanguage to derive the understanding of the object language.
According to Dummett, theories of meaning can further have the following characteristics:

- In a *rich* theory of meaning the knowledge of the concepts is achieved by knowing the features of the concepts. An *austere* theory merely relies upon simple recognition of the shape of the concepts. A rich theory can justify the use of a concept by means of the characteristic features of that concept, whereas an austere theory can justify the use of a concept merely by enumerating all occurrences of the use of that concept.
- A *molecular* theory of meaning derives the understanding of an expression from a finite number of axioms. A *holistic* theory, in contrast, derives the understanding of an expression through its distinction to all other expressions in that language. A molecular theory, therefore, gives criteria to associate a certain meaning to a sentence and can explain the concepts used in the language. In a holistic theory nothing is specified about the knowledge of the language other than in global constraints related to the language as a whole.

A *modest holistic* theory is explicit about the knowledge of meaning because it can enumerate all occurrences of a certain proposition, but it does not model what constitutes the knowledge of that proposition. Dummett accordingly concludes that *a holistic view of language renders the construction of a systematic theory of meaning impossible.*[p. 123]
On the other hand a *modest molecular* theory of meaning cannot explain what the knowledge of the propositions consists of other than in an understanding of one language via an understanding of another because it presupposes an understanding of the metalanguage in which the propositions are expressed. This, he argues, is just what a translation manual does.
A *full blooded* theory of meaning has to explain what someone knows if he knows the language. It must explain why and when a concept can be applied correctly and — in case of dispute — it must be able to justify the use of a concept by giving the respective features of the concept. A full blooded theory, therefore, needs to be molecular and rich and it needs to be described in a language that does not assume any of the primitives to be already known.
Coming back to the Machine-Translation issue, it thus appears that all theories of meaning are equivalent to a translation manual. In order for the Evaluation process to communicate the set of rules (which describes the theory) to the system, it must use a language, and in order for the system to execute the rules it must have an operative knowledge of the primitives used in that language to perform the expected input/output mapping. This conclusion could be avoided if the Evaluation process communicates the mapping "rules" not by means of a language but rather by means of, say, a set of unstructured patterns. But, surely, such a set could hardly count as the description of a theory.

The theory of meaning to be communicated to the system in the OMT-scenario is equivalent to a translation manual because 1) a full blooded theory of meaning is impossible to formulate, 2) a modest holistic theory is uninteresting because it is necessarily unsystematic and 3) a modest molecular theory of meaning is just what a translation manual does.
In the next section I will therefore examine *epistemological* MT-systems which implement translation manuals in a direct manner.

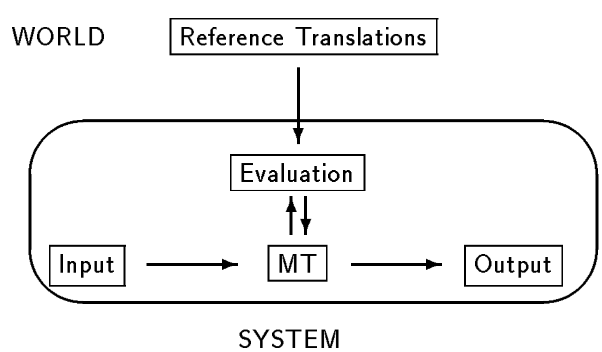
Epistemological Machine Translation

Epistemological systems change their 'rules' according to the needs of the input/output mapping. Because the Evaluation process is part of the system itself, epistemological systems are second order systems that learn the meaning preserving mechanism of the input/output relation. (i.e. understand the mapping requirements according to the system-external WORLD).
In order for such systems to learn the Evaluation process, a learning corpus is given. This corpus constitutes the reference of the Evaluation competence. A translation $T_{source} \leftrightarrow T_{target}$ is thus true for the system if and only if an appropriate reference translation $R_{source} \leftrightarrow R_{target}$ is available in the corpus. Note that the above equivalence (1) turns into (2):

$$T_{source} \leftrightarrow T_{target} \Longleftrightarrow R_{source} \leftrightarrow R_{target} \qquad (2)$$

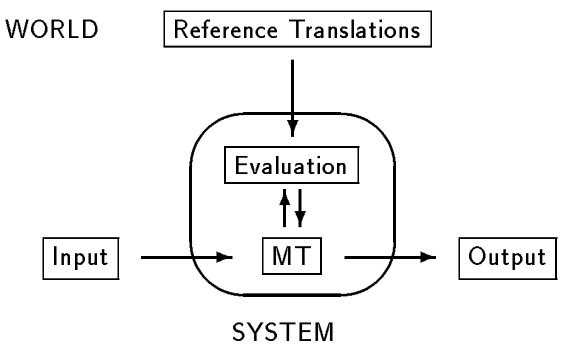
Epistemological MT-systems (EMT) implement a translation manual in a direct way: the understanding of the appropriate translation mapping ($T_{source} \leftrightarrow T_{target}$) is derived from the understanding of the reference translation mapping ($R_{source} \leftrightarrow R_{target}$).
The interesting question in EMT is: how is the Evaluation process designed that generates an appropriate understanding of the mapping mechanism? Similar to OMT-2, the EMT-1 scenario as shown in Figure 3 translates a restricted type of text. The difference between EMT-1 and OMT-2 is that the former derives the appropriate set of rules from the learning corpus, while the latter receives these rules from the outside WORLD. However, both are similar in the sense that they are designed for a certain restricted type of texts.

Figure 3
EMT-1: The MT-system in the oval frame translates restricted input text. The Evaluation *is part of the system and performs translations according to the references given.*



If it were possible to give a representative corpus of all translations for a source language and a target language one might hope to generalize EMT-1 to become an all-purpose MT-system. The Evaluation process in an EMT-2 scenario in Figure 4 would thus have sufficient references to perform an all-purpose MT just as it is intended in OMT-1.

Figure 4
EMT-2: The MT-system in the oval frame translates unrestricted input text. The Evaluation *has sufficient references to perform all-purpose translations.*



Unfortunately, epistemological systems can only be realised through non-trivial machines that are far too complex to be computed as I will discuss in the next section.

Why EMT-2 cannot work

According to [Foe93], trivial machines are characterized by a fixed machine table: the same input symbols always triggers the same (set of) transformation rules that generate the same output symbols. Transformation rules in trivial machines do not change, they are independent from past experiences; the outcome is predictable because it only depends on the input symbols and the (fixed) set of transformation rules. Ontological systems are such trivial machines. Once a set of rules given, the system always generates the same output for a given input.
In contrast to trivial machines, non-trivial machines change the machine table according to a "program of second order" (i.e. the Evaluation process). The generator of the transformation rules (the Evaluation process) is part of the system. Non-trivial machines are thus capable of adapting to a changing environment because they have a learning component. However, non-trivial machines are, generally, far too complex to be analytically determined because the number of possible machines is far too big to be computed.
According to [Foe93], the number of non-trivial machines $M_S(X,Y)$ which have S internal states, X input symbols and Y output symbols is $M_S(X,Y) = Y^{SX}$. Thus, if the number of input symbols and output symbols is $X = Y = 4$ and the number of internal states is $S = 24$ then the number of possible machines is $M_{24} = 4^{24\times 4} \sim 6.3\times 10^{57}$. Even in such a small setting it is impossible to determine a concrete machine because even if we assume that one Million machines can be checked per second, the universe is still too young to check only a small part of them.

If, now, we assume an MT-system to consists of 5 rules (i.e. the number of internal states is $S = 5$) and if we assume the number of input and output symbols (i.e. morphemes in the respective source and target language) to be at least 10.000 ($X = Y = 10.000$), then according to the above formula there are $10^{200.000}$ different machines among which we need to find the one that implements the all-purpose MT-system.
In a more realistic application, if we exclude "free" translation by restricting possible output symbols for a given input symbol to 3 ($Y = 3$), if we restrict the maximum sentence length to 20 (i.e. $X = 20$) and if we assume the same underlying system, we still have $3^{100} \sim 5.1 \times 10^{47}$ possible machines.
However, for arbitrary input text, this seems too strong a restriction because there are many sentences that contain more than 20 words and often more than one sentence needs to be considered to arrive at a valuable translation. Furthermore, one cannot be certain that 5 rules (i.e. 5 internal states) are sufficient to process arbitrary input text.
Therefore, the best we can hope for is to approximate valuable translations for a restricted domain according to a scenario as illustrated in EMT-1 and OMT-2. The emerging research interest in the area of controlled language is a consequence of this: one tries to trivialize language in order to make it ready to be processed in MT-systems (and other NLP-applications).

Approaches to EMT-1

Several MT paradigms are subsumed under the EMT-1 approache. These systems have in common that, given the reference corpus, first an appropriate set of mapping rules is generated according to the 'understanding' capacities of the Evaluation process. In the translation phase, these mapping rules, then, are responsible for the input/output mapping.
According to the above classification one can distinguish between *holistic* approaches, *austere* approaches and *rich* approaches.
Statistics based MT approaches e.g. [BCDP+90] have a *holistic* view on languages. Every sentence of one language is considered to be a possible translation of any sentence in the other language. In order to compute the most probable translations, each pair of items of the source language and the target language is associated with a certain probability. This a priori probability is derived from the reference corpus. In the translation phase, several target language sequence are considered and the one with the highest (posterior) probability is then taken to be the translation of the source language string.
In such a system, no account is given for the equivalence of the source language meaning and the target language meaning other than by means of global considerations concerning co-occurrence frequencies in the reference corpus. Although one can imagine having rich holistic MT-systems, to my knowledge only austere holistic approaches have been investigated. In rich holistic MT-systems the translation probability would not only rely on the shape of the words but rather on their (linguistic) properties.

However, good results have been reported for austere holistic systems if the learning corpus contains several million translation examples.

Translation memories (TM) (e.g. TRADOS [Hey96], TRANSIT) represent *austere* approaches to MT. A typical TM relies solely on the similarity of the shape of the source text and the reference corpus. The target language equivalents of the most similar candidates are then presented as the translation of the source text.

TMs are molecular because they can display the difference in the source text and in the retrieved reference text. They can enumerate all occurrences in the reference corpus that contain the use of a certain concept. They cannot, however, justify the use of a word other than by enumerating all contexts in which the word occurs. Because these systems are easy to configure and quickly adaptable to different types of texts, TMs represent one of the most popular approaches to epistemological MT.

Example Based Machine Translation systems (EBMT) have a *rich* view on languages. In [CC97] morphological analysis and syntactic chunking of the reference corpus is carried out. Abstract templates are generated that contain variables in those positions where the source language and the target language equivalences are strong. In the translation phase, a multi-layered mapping from the source language to the target language is processed on the level of templates and on the level of fillers.

Such systems are molecular and rich because the mapping rules function in a compositional manner. According to the way in which abstractions are derived, they can justify and give account in what consists the similarity of two (or more) concepts.

Conclusion

It is widely acknowledged that translations from one language into another are valid iff the source language text and the translated target language text have the same meaning. In this paper I have investigated how this meaning preservation requirement is handled in MT-systems.

Two approaches to MT are distinguished: in *ontological* MT, a set of rules is given to the system from an outside Evaluation process that enables the system to map an input text onto an appropriate (i.e. meaning preserving) output text.

In contrast to ontological MT-systems, *epistemological* MT systems induce the set of mapping rules based on a set of translation examples (a reference corpus).

Whereas for ontological MT-systems the meaning of a text is captured by a set of rules that implements a theory of meaning, epistemological MT-systems induce an appropriate theory and generate accordingly a set of mapping rules based on a given corpus of examples.

Theories of meaning may have the following characteristics: an *austere* theory of meaning relies merely on the recognition of the shape of the text (and the words it contains), whereas a *rich* theory of meaning 'knows' the constituting features of the concepts used in the language.

A *holistic* theory of meaning derives the meaning of an expression through its contrastive use in a language, whereas *molecular* theories derive the meaning of an expression from finitely many axioms.

In the light of these characteristics different approaches to epistemological MT are discussed: Statistical MT represent a holistic view on languages, translation memories realize an austere approach and example based machine translation follows the rich approach. However, none of the systems is capable of implementing all-purpose MT.

Acknowledgement

I whish to thank Oliver Streiter, Cath Pease, Antje Schmidt-Wigger and Johann Haller for fruitful discussions and comments on former versions of this paper.

References

[BCDP+90] Peter F. Brown, J. Cocke, Stephen A. Della Pietra, Vincent J. Della Pietra, F. Jelinek, Robert L. Mercer, and P.S. Roossin. A statistical approach to machine translation. *Computational Linguistics*, 16:79–85, 1990.

[CC97] Bróna Collins and Pádraig Cunningham. Adaptation guided retrieval: Approaching EBMT with caution. In *TMI-97*, pages 119–127, 1997.

[CG96] John Chandioux and Annette Grimaila. "Specialized" Machine Translation. In *Proceedings of the Second Conference of the Association for Machine Translation in the Americas*, pages 206–212, Montreal, Canada, 1996.

[Dum75] Michael Dummett. What is a Theory of Meaning? In *Mind and Language*. Oxford University Press, Oxford, 1975.

[Foe93] Heinz von Foerster. *KybernEthik*. Merve Verlag, Berlin, 1993.

[Fre62] Gottlob Frege. Sinn und Bedeutung. In *Funktion, Begriff, Bedeutung*. G. Patzig Vendenhoek & Ruprecht, Göttingen, 1962.

[Hey96] Matthias Heyn. Integrating machine translation into translation memory systems. In *European Asscociation for Machine Translation - Workshop Proceedings*, pages 111–123, ISSCO, Geneva, 1996.

[Put75] Hilary Putnam. *The Meaning of 'Meaning'*. University of Minnesota, Minneapolis, USA, 1975.

[Sch94] J. Siegfried Schmidt. *Kognitive Autonomie und Soziale Orientierung*. Surkamp, Frankfurt, 1994.

[Tar35] A. Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. In *Logik-Texte*, pages 443–546. Wissenschaftliche Buchgesellschaft, Darmstadt, 1935.