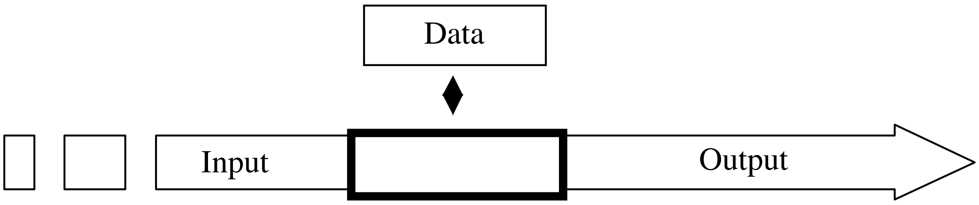
Knowledge Representation in Machine Translation

Walther v.Hahn
Computer Science Department, University of Hamburg

Summary:

This paper describes the use of declarative non-linguistic knowledge in machine translation systems. Different types of knowledge and their use in three typical system environments are discussed. Special emphasis is put on the benefits of knowledge bases in others than interlingua systems. Finally, the perspectives for using ontologies emerging from the "Semantic Web" activities are sketched.

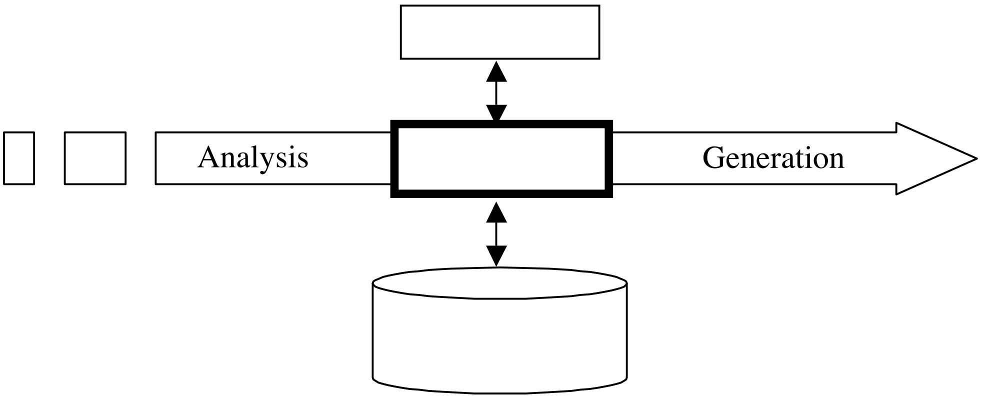
The history of knowledge representation – in its very basic form – starts, when programmers distinguish between their program and the data which the program processes,



i.e. when lexicons, grammar rules etc. were not any longer part of the program itself, written in the code here and there, but modularized out of the program and read from a file whenever the program is executed. This distinction assumes already that there are separable data (something like "tree" = noun, e.g.) in a program.

The reader may argue, that every program applies the programmer's knowledge by being executed, thus being a knowledge based program. But we don't go that deep into programming theory and simply assume, that normally the program code itself does not include the lexicon and the grammar rules. This declarative knowledge (in contrast to the procedural knowledge of the program, is usually written in separate files and used by program calls.

However, talking about knowledge in machine translation, refers to the *additional* use of non-linguistic knowledge [CMUZZ96], which can be accessed for semantic disambiguation, word sense relations, inferences, etc.:



According to the place, where this additional type of knowledge is integrated in machine translation, research distinguishes three types of systems [HuSo92]:

1. Systems using terminological material, which is systematically ordered along a schema of the (technical) field. These systems, however, do not contain explicit, i.e. declarative knowledge bases of their domains.

2. Systems using knowledge about concepts or facts for specific tasks like syntactic disambiguation, word sense disambiguation, or pronoun insertion.

3. Systems that construct a deep meaning representation (in most cases interlingua systems) by using additional knowledge of some sort.

Non-linguistic knowledge in the sense of systems (2) and (3) consists of three types:

a) Conceptual knowledge, which describes the top-level of an ontology, relations among concepts and general inferences about concepts

□b Bear (b) □ Animal (b)	"Bears are animals"
□b Bear (b) □ SpeciesOf(b)=Ursidae	"Bears belong to the species of Ursidae
□x,y part of (x,y)& PhysicalThing (y) □ PhysicalThing (x)	"All parts of physical things are physical things"

b) World knowledge and facts, which may include measures, time, space, events, etc.

□x PhysThing (x) □ □s Size(x) = s	"All physical things have a size"
Bear (Pooh)	"Pooh is a bear"
T(Area(Poland, SqMiles(233000)),AD1426)	"In 1426 Poland had the size of 239000 square meters"

and

c) Situation knowledge, which describes the situation in which the text is situated:

In (Pooh, LivingRoom$_3$)	"Pooh is in the living room"
Female (Speaker)	"The current speaker is female"

(Examples from Russell/Norvig [Rus95])

Whatever a translation system does, it needs at least some types of this knowledge. If the functionality is not very demanding, some parts of the conceptual knowledge can be treated in the lexicon, e.g. under semantic subcategorization, roles or constraints.
The following examples may show that each of the knowledge types are needed even if the translation is done by a transfer architecture:

- Lexical disambiguation (source: facts):

"I **go** to my office in an hour" □ "Ich **gehe** in einer Stunde in mein Büro"
□ "Ich **fahre** in einer Stunde in mein Büro"
□ "Ich **fliege** in einer Stunde in mein Büro"
(engl.□ germ.)

dependent on how far away my office is.

- The previous example becomes even more complicated if the semantics of the prepositions is included (source: concepts/facts):
“I read a story about evolution **in** the last million years”
 “Ich las einen Artikel über die Evolution **während** der letzten 1000 Jahre”
“I read a story about evolution **in** ten minutes
 “Ich las einen Artikel über die Evolution **in** 10 Minuten”
(engl. germ.)

- Anaphora antecedents (source: concepts/facts):
“She took the ice cream out of the fridge and ate **it**” Sie nahm das Eis aus dem Kühlschrank und aß **es**” (not “**ihn**”) (engl. germ.)

- Explicit Pronouns (source: facts):
“**Sie** gehen” **Ei** merg / **Ele** merg / **Ei si ele** merg (germ. romanian)
Dependent on whether the set of persons (“Sie”) contains men or woman or both

- Lexical de-specification (source: concept hierarchy):
“Cousine” /” Cousin” “cousin” (germ. engl.)

- Lexical specification (source: concept hierarchy):
“Onkel“ "farbror” or “morbror” (germ. danish)

Example I

The KBMT project [KBMT89] was the first systematic attempt (completed 1989) to use knowledge representation for a deep representation of the contents of source language sentences (an interlingua).

The assumption behind KBMT [NiCar92] is that

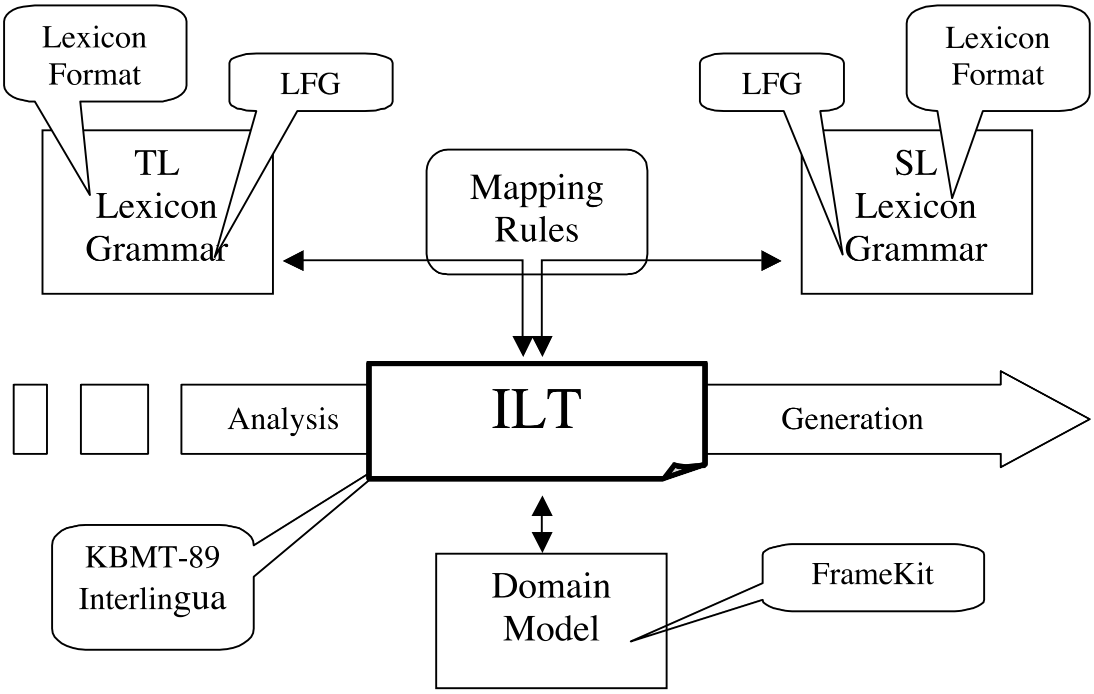
a) One “functionally complete” meaning representation can serve for translations to a number of languages, and
b) no total representation of human understanding of a text is necessary.

The KBMT schema is intended to function in domains which are relatively unambiguous, e.g. technical documents. Representing the complete knowledge about rather open subjects (like cultural events) is practically impossible.

Basic components of a KBMT system are:

- An ontology of concepts (“domain model”)
- Source language (SL) lexicon and grammar for the analysis process
- Target language (TL) lexicon and grammar for the generation processes
- Mapping rules between the Interlingua and SL/TL syntax.

The KBMT-89 system delivers bidirectional translations for English and Japanese and is designed for translation of PC manuals. As input KBMT-89 accepts single sentences of English or Japanese; in the analysis step their meaning is represented as “interlingua text" (ILT). The KBMT-89 generator eventually processes the ILT and produces Japanese or English sentences, respectively.



For these steps the following formalisms are used:

1. A knowledge representation system FRAMEKIT, based on frames,
2. A LFG-like grammatical representation language
3. A formal language specially constructed for representing text meanings (the "interlingua" language);
4. The lexicon formalisms with structural mapping rules

A FRAMEKIT frame consists of named slots of features and values, where values may be frames again. For example the following frame illustrates a possible meaning representation for the sentence *"Save the document"* [Tru99]*:*

$$\begin{bmatrix} instance\_of: & save \\ isa: & physical\_event \\ id: & save\_43 \\ agent: & user \\ patient: & \begin{bmatrix} instance\_of & document \\ isa & separable\_entity \\ id: & document\_72 \\ reference: & definite \end{bmatrix} \end{bmatrix}$$

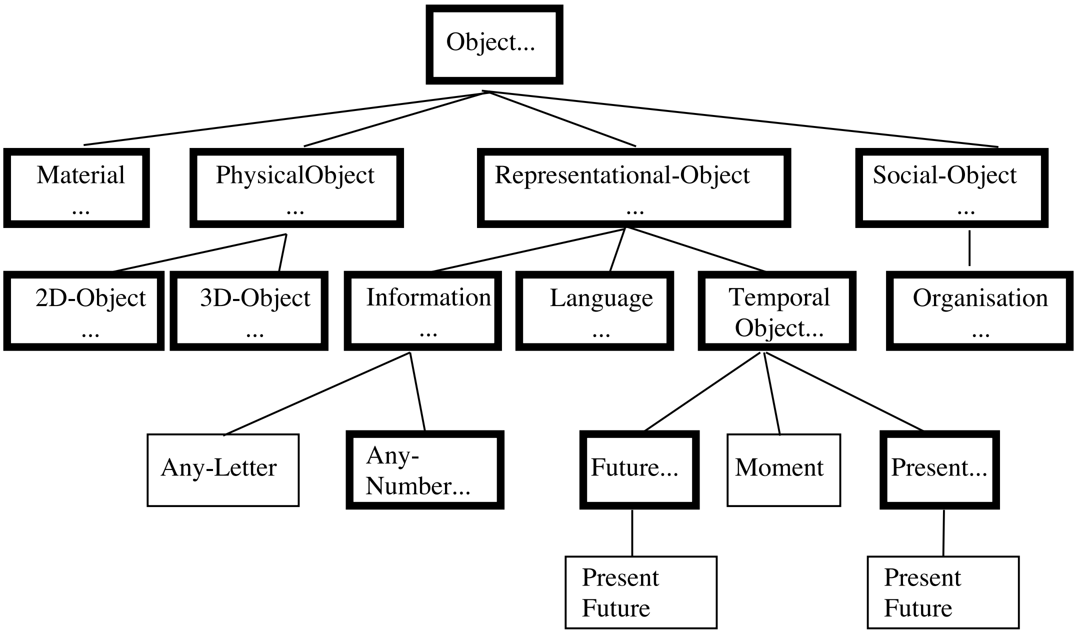
This example shows that frames contain linguistic and non-linguistic knowledge. A powerful feature of frames is their support of inheritance mechanisms: For example, a frame PERSONAL_COMPUTER with all general information about PCs is already

defined. A new frame for a specific PC brand does not need to repeat all this information, but represents only its specific features, everything else is inherited along the isa-slot.
The KBMT-89 ontology is a language-independent conceptual representation of the interaction between personal computers and their users.

The KBMT-89 ontology contains:

```
Objects
Events
Properties of objects or events
Relations
Attributes
```

Concepts are linked to others by relations. Each concept has attributes which specify value sets. Value sets contain only literals (i.e. no concepts):



Fragment of the ontology of KBMT-89
(Concepts in bold boxes = further sub-concepts are omitted for clarity)

The concept lexicon of the system consists of the domain ontology plus the lexical mapping rules. The concept lexicon assigns semantics to all lexical items of a sentence by using lexical mapping rules.

Below an example of the parser’s output for the sentence: *“Get the diagnostics diskette from the back of this manual”*

```
[*RECEIVE
   (AGENT *READER)
  (THEME [*DISKETTE
                (NUMBER SINGULAR)
                (PURPOSE [*DIAGNOSE
                                 (NUMBER SINGULAR)])
                                 (REFERENCE DEFINITE)])
                (SOURCE [*BACK-OF-3D
```

```
               (NUMBER SINGULAR)
               (REFERENCE DEFINITE)
               (PART-OF [*MANUAL
                          (NUMBER SINGULAR)
                          (REFERENCE DEFINITE)
                          (NEAR +ü)])])
(TENSE PRESENT)
(MOOD IMPERATIVE)
(CLAUSAL-MARK +)
(NUMBER-BULLET [*ANY-NUMBER
                (CARDINALITY 1)])]
```

From the parser's output a process called Augmentor produces the ILT.

As mentioned, even systems that do not follow this interlingua approach, will use non-linguistic knowledge in one or another form. A good example is the speech-to-speech translation system Verbmobil, which uses, e.g., knowledge in a KL-ONE like format [NoHaf97] for contextual disambiguation. Furthermore, domain-dependent dialog act schemata are used to reduce the search space for translation.

Example II

The system **Verbmobil** [vHaTes00][Wah00] is a transfer-type system between English, Japanese and German. It translates spontaneous speech input on-line and near real-time. The central representation structures are VITs (Verbmobil Interface Terms) for each input, which represent the necessary linguistic and non-linguistic information. The following expression is the representation of the input *“He is coming at the beginning of August”*. The comments (%) indicate the representation layers. It can be seen easily that VITs do not contain a representation of the semantic contents or the pragmatic sense of an utterance. A unique feature of VITs, however, is their representation of dialogue acts (similar to speech acts, but domain action dependent), and the use of prosody as indicator for structural boundaries, particle interpretation and sentence mood.

```
Vit( vitID(sid(104,a,en,10,800,1,en,y,semantics), %SegmentI
      [word(he,1,[l126]),                            %WHG string
         word(is,2,[]),
          word(coming,3,[l127]),
          word(at,4,[l136]),
          word(the,5,[l128]),
          word(beginning,6,[l135]),
          word(of,7,[l135]),
          word(“August”,8,[l134])]),
      index(l138,l125,i35),                         %Index)
      [beginning(l135,i37),                         %Conditions
       arg3(l135,i37,i38),
       come(l127,i35),
       arg1(l127,i35,i36),
       decl(l137,h43),
       pron(l126,i36),
       at(l136,i35,i37),
       mofy(l134,i38,aug),
       def(l128,i37,h42,h41),
```

```
       udef(l131,i38,h45,h44)],
      [in_g(l126,l125), in_g(l137,l138),          %Constraints
       in_g(l127,l125), in_g(l128,l130),
       in_g(l131,l133), in_g(l134,l132),
       in_g(l135,l129), in_g(l136,l125),
       leq(l125,h41), leq(l125,h43),
       leq(l129,h42), leq(l129,h44),
       leq(l130,h43), leq(l132,h45),
       leq(l133,h43)],
      [s_sort(i35,situation),                     %Sorts
       s_sort(i37,time),
       s_sort(i38,time),
      [dialog_act(l125,inform),                   %Discourse
       dir(l136,no),
       prontypel136,third,std()],
      [cas(i36,nom),                              %Syntax
       gend(i36,masc),
       num(i36,sg), num(i37,sg), num(i38,sg),
       pcase(l135,i39,of)],
      [ta_aspect(i35,progr),                     %Tense and Aspect
       ta_mood(i35,ind),
       ta_perf(i35,nonperf),
       ta_tense(i35,pres),]
      [pros_accent(l135,progr)],                  %Prosody
)
```

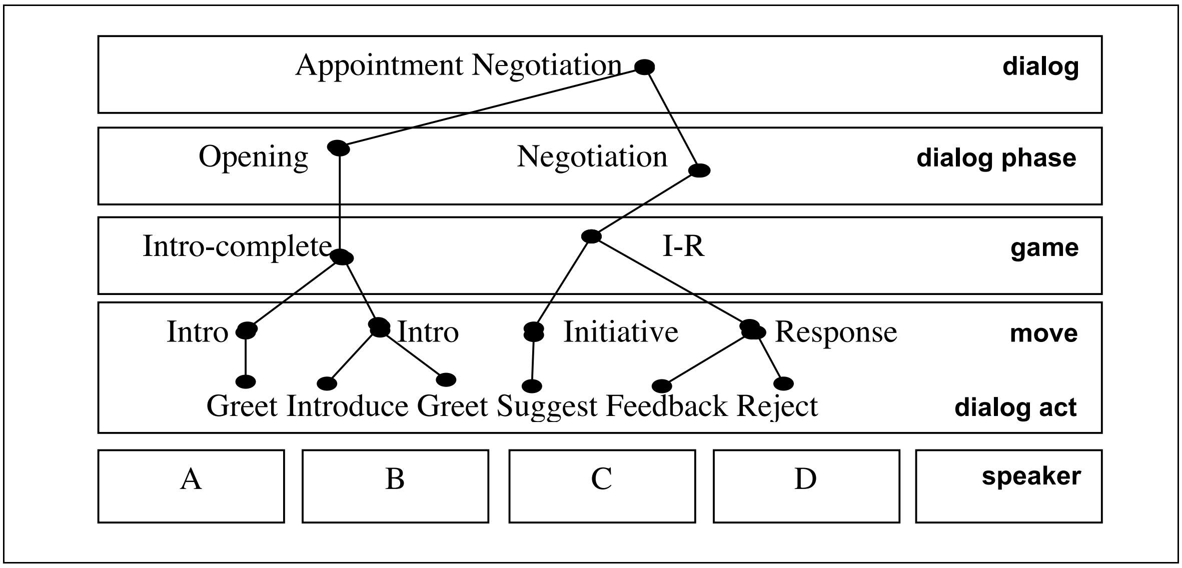
The construction of VITs relies, among others, on a multilingual semantic data base (SemDB), which contains:

- Base form of words,
- Lexical semantic decomposition,
- Interlingual representation, if available,
- Semantic class (nominals, quantifiers, verbs, modifiers, etc.),
- Syntactic valency with mappings onto grammatical functions and thematic roles (linking),
- Ontological sorts, e.g.:

abstract
        property, field, info-content, institution, symbol
    space-time
        temporal
            situation
                meeting_sit, communicate_sit, action_sit, …)
            time
        entity
            object
                agentive, thing
            location

- Selectional restrictions on arguments.

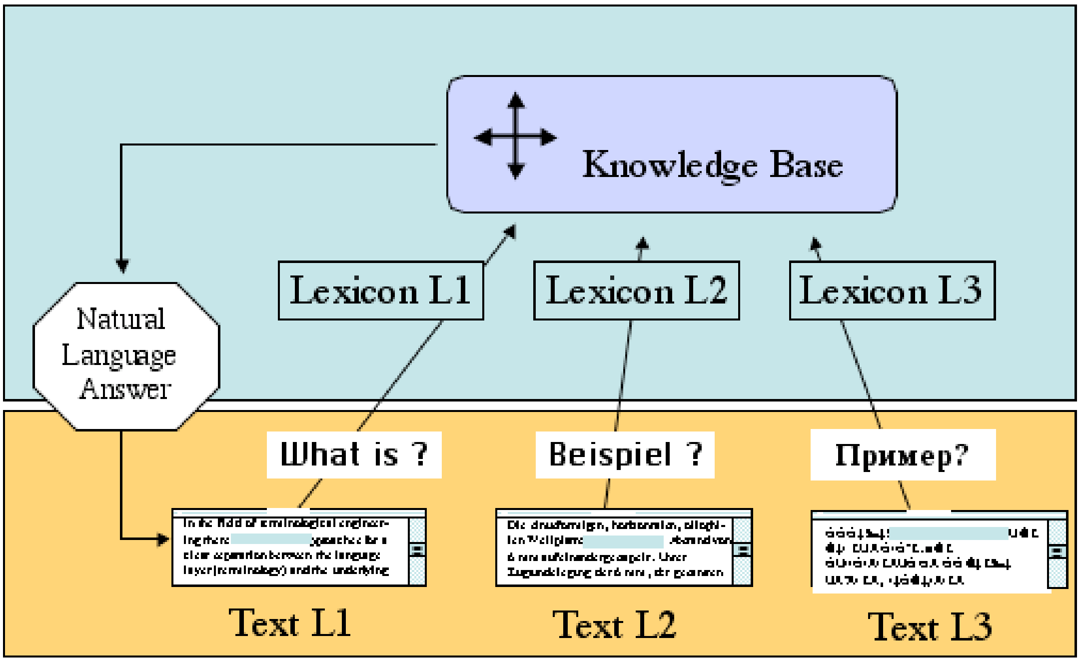
The dialog act representation is used for syntactic disambiguation and repair of gaps in the set of speech recognition hypotheses, e.g.. The following schema shows the four levels [vHaJek00] of a negotiation dialogue:



The use of knowledge is not restricted to automatic translation systems. Machine aided translation also benefits from general knowledge as can be shown in the system DBR-MAT [vHaAn94], a system to support technical translators. It uses elementary domain knowledge to explain the conceptual background of an utterance. This may be interesting for non-specialized translators in not frequently used languages.

Example III

In DBR-MAT, the translator can ask conceptual questions, e.g. nested intensional and extensional definitions about basic domain knowledge, questions about characteristics, hyponyms or meronymies [vHa97].



The background representation is an ontology and a broad conceptual description, written in “Conceptual Graphs” [Sowa84], enriched by linked pictures. Target and source language are linked separately to the semantic representation, because everybody knows that the lexicalization of languages is different [AnBo96b].

The following expressions describe two complex conceptual relations (“situations”) in the domain of environmental technology:

```
[SITUATION: [OIL FRAGMENT: {*}] □  (IN) □  [WATER: {*}]
        □  (CHAR) □  [PHYSICAL STATE:
           disj{MEMBRANE, DROPS, COLLOID, EMULSION, SOLUTION}].

[SITUATION:
        [WASTE WATER: {*}] □ (CONTAIN) □ [OIL FRAGMENT: {*}]
             □ (ATTR) □ [FLOATING]
             □ (ATTR) □ [ROUGHLY DISPERSED]
       □ (PTNT) □ [PRECIPITATION]].
```

The knowledge base of DBR-MAT contains the following objects [AnBo96a]:

KB Objects	Example
Concepts	[OIL SEPARATOR]
Individuals	the separator C334
Conceptual relations	(part_of)
Contexts (situations)	[SITUATION: [WASTE WATER: {*}] -> …
A type hierarchy	[SEPARATOR] ↓ [OIL SEPARATOR]

Submenu	***Item***	***Evaluated Conceptual Relations***	***Inheritance***
What is?		Types of… - All + ATTR, Char, PART_OF	✓
Types of -	All	Superconcepts + subconcepts + sister concepts	
	General	All superconcepts from the hierarchy	
	Concrete	All subconcepts from the hierarchy	
	Similar	All sister concepts from the hierarchy	
Characteristics -	All	Attributes + Who + Object + How + Where	✓
	Attributes	ATTR + CHAR	✓
	Who	AGNT	
	Object	OBJ + PTNT	
	How	INST	
	Where	LOC + DEST + FROM + IN + TO	
More…		All remaining relations	
Examples		Individual concepts	
Want All		All mentioned above, without duplicates	✓

The table gives

1. a list of all conceptual relations, that are used in the representation (in the third column) and
2. the rules, how DBR-MAT reads the knowledge base, triggered by specific questions (column 1).

The three examples demonstrate the need for conceptual knowledge and world knowledge in different system types and architectures.

Coming back to distinctions mentioned in the beginning of the paper: Systems need non-linguistic knowledge to solve a number of linguistic tasks. The way, however, how explicit, how declarative and how modular the knowledge is represented distinguishes classes of systems.

Verbmobil, e.g., has a very modular (and partly parallel) structure and follows as far as possible a declarative paradigm with widely accepted formalisms. DBR-MAT uses a standard representation language and, e.g., a declarative editable table of rules for traversing these CG-graphs (see above), to ensure an open inferencing behaviour of the system, instead of integrating the rules into the traversing program. In KBMT-89 most knowledge bases are modular and declarative, but at that time no widely accepted standards for interchange could be applied.

Many other translation systems have significant problems even with basic conceptual knowledge, because it is hidden somewhere in processes or is included in semantic features, subcategorizations or case roles. In such systems the designer never knows, how much knowledge is represented in total, and where. Additionally, there is no way to check multiple representation, because most of the procedural knowledge is local. Declarative knowledge sources, in contrast, can be maintained in isolation, can be exchanged and may be used in/from other inference machines or grammars. It even can be used in other systems than translation systems. Esp., problems (1) and (2) of the following list increase with implicit and procedural representations.

Critical problems of knowledge-based systems are still

1. The effort to build up knowledge bases,
2. A practical definition of the size (or the coverage) of the knowledge base,
3. The choice of an adequate granularity of the knowledge, and
4. The choice of an adequate representation language and its necessary logical/formal properties.

The basic idea of declarative, modularized and system independent knowledge has recently become very important for the development of the Internet after the famous papers of Berners-Lee in 2001 [BernLee01] on the concept of the „Semantic Web“. The Semantic Web (according to the definition of the W3C consorium) is the abstract representation of reference concepts (a subset of the domain knowledge) on the World wide Web, based on the RDF standards and other standards to be still defined. The Semantic Web is being developed by the W3C, in collaboration with a large number of researchers and industrial partners from various communities as: computer science, computational linguistics, mathematics, logics, knowledge management, e-commerce etc.

[BernLee03] claims that

> *"The semantic web will facilitate the development of automated methods for helping users to understand the content produced by those in other scientific disciplines. On the semantic web, one will be able to produce machine-readable content that will provide, say, automated translation between the output of a scientific device and the input of a data mining package used in some other discipline, or a self-evolving translator that allows one group of scientists to directly interact with the technical data produced by another.*
>
> *These new products will allow users to create relationships that allow communication when the commonality of concept has not (yet) led to a commonality of terms. The semantic web will provide unifying underlying technologies to allow these concepts to be progressively linked into a universal web of knowledge, and will therefore help to break down the walls erected by lack of communication and allow researchers to find and understand products from other scientific disciplines".*

Briefly, the semantic web will be based on an ontology (or several ontologies) per domain to which all WWW sites can refer to express the semantics of their linguistic and non-linguistic objects. The *Semantic Web*, consequently, can then be used

> for *semantic* web operations like data mining, information extraction, summarisation etc., because there exists one semantic reference, which is agreed upon among the users of the net, and by which the URL receives its contingency,

where nowadays standard *information technologies*

> can only retrieve (groups of) words, collect statistics about there occurrence or co-occurence of words and extract metainformation from the header of web pages. In the best case the corresponding field of the web site has a terminology, a conceptual taxonomy or a nomenclature (e.g. in medicine), which might be used for semantic extraction. But not every URL in every field will refer explicitely to such an existing knowledge rrepresentation system.

According to [Benjamins&02], one of the big challenges for the Semantic Web is multilinguality. The Semantic Web is composed of concepts, relations among concepts and logical rules holding for these relations (like "transitivity"). The concepts have abstract (maybe English) names (designators), but they are not words by themselves. The names are arbitrary and unique attributes of the concept and thus have no ambiguity or other features of natural language words. The semantics of a concept is defined by its position in the ontology, e.g. by being a subconcept or superconcept af another concept. The abovementioned rules may define that, e.g, if B is a superconcept of C and A is a suberconcept of B, than A is also a superconcept of C.

Even though English is the predominant presentation language in WWW, there exist important Web resources in other languages: Japanese 5.9%, German 5.8%, French 3.0% etc. These languages will use the Semantic Web, too, and pages in German, French, Japanese etc. will be utilized by users of other languages. Today we may use (or not) web translation services, however, they do not have enough information about the semantics of retrieved pages to deliver translations of sufficient quality. The Semantic Web my overcome such problems

Thus, there are two main future relations among SemanticWeb and Machine Translation:

1) Support of Machine Translation from the Semantic Web

In particular, for machine translation the result of the semantic web activities will be that widely accepted ontologies of domains will be available as large standardized knowledge bases, to which terminology from a specific natural language can be attached.. This will solve at least items (1) and (3) out of the above mentioned list of problems.

To sketch a very simplified example: The semantic web ontologies can support the choice of hyponyms or hyperonyms for lexical specification or de-specification (see above) by giving the superconcepts or subconcepts in the domain at hand.

2) (Machine) Translation for the Semantic Web

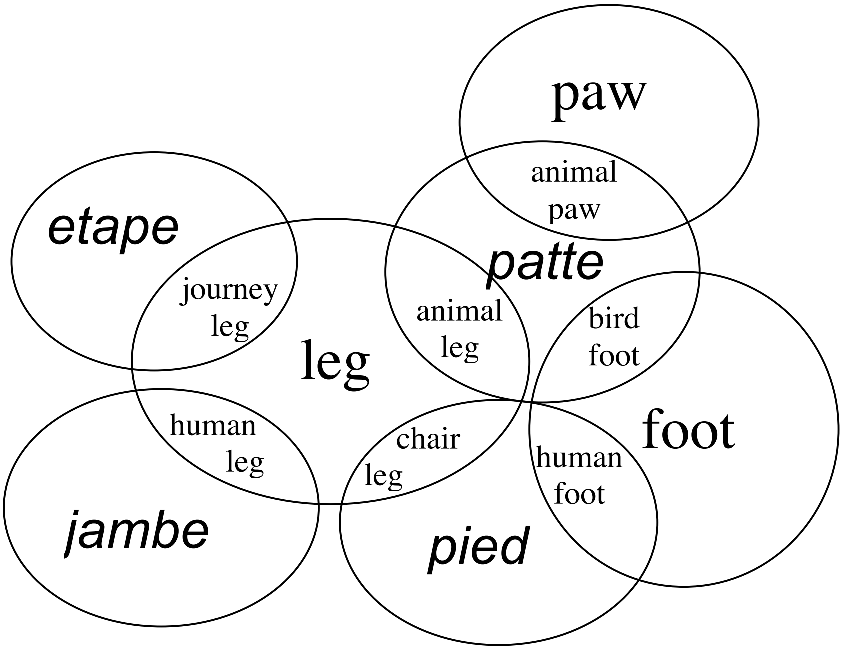
The development of ontologies and the annotation of web resources for the Semantic Web raise problems both on ontology level and annotation level.

> In the development of ontologies it is desirable, that concepts have natural language names attached to them, i.e., words of an existing language, which can be used for building definitions or retrieve non-English texts.
>
> Annotation of Web objects (reference to ontologies) is a process where a large amount of users (content providers) are involved. The annotation (via the format RDF, e.g.) is the essential step for the semantic use of web ressources. Therefore it is desirable to provide RDFS repositories for as many languages as possible. The objects in the repository can afterwards be mapped onto one and the same international ontology.

Earlier pilot-projects projects like PANGLOSS [KnightLuk94] showed that for unrestricted domain this mapping raises problems. PANGLOSS aimed at constructing a large ontology for supporting KBMT. The goal was to scale KBMT up from specific constrained domains to all newspaper texts. The ontology was constructed by merging various online dictionaries (Longman), semantic networks (WordNet), PENNMAN upper model), and bilingual resources (Spanish-English Collins) through semi-automatic methods (conceptual matching of semantic taxonomies, bilingual matching). The main problems raised with merging bilingual resources, because the tools tried to map bilingual linguistic entries onto the same ontology. Partially overlapping words, words with different degrees of polysemy, with independent domain senses, words with different stylistic or historical range, etc. made it impossible to solve this task once and for all domains.

Example: the Spanish word “manzana” can be translated as block in English but maps only one of the concepts referred by block, namely CITY-BLOCK but does not map with BUILDING-BLOCK.

The complex overlap between English *leg, foot and paw* and various French translations is a good example [JurafMartin00]:



The examples above demonstrate rather clearly that a one-to-one mapping between annotation objects, words (even not expressions) in different languages and concept names in RDFS repositories is not possible. The mapping between an entry in an RDF repository and an ontology concept has to take into condsideration three dimensions: the lexical dimension and its similarities, the domain and the closeness of its concepts, and the language pair with the translation relations.

The linking between these objects within an ontology is currently tried with the help of the DAML-OIL Standard which for the moment provides only relations as: sameClasssAs and samePropertyAs to refer to identical corresponding classes or properties in different natural languages.

In the future, experience from the development of other multilingual databases (e.g. EuroWordnet [Vossen98]) has to be used to provide flexible relations like “Synonym”, or “Near_Synonym”, which are used there.

References:

[AnBo96a] Angelova, G. and K. Boncheva. *NL Domain Explanations in Knowledge Based MAT*. In: Proc. COLING-96, Copenhagen, Denmark, pp. 1016 - 1019.

[AnBo96b] Angelova, G. and K. Boncheva. DB-MAT: a NL Based Interface to Domain Knowledge. In: A. Ramsay (Ed.), Proc. AIMSA-96, IOS Press, Vol. 35 in the series ”Frontiers in AI and Applications”, pp. 218 - 227.

[Benjamins&02] Richard Benjamins, Jesús Contreras, Oscar corcho,Asunción Gómez-Pérez, „Six challenges for the Semantic Web“ . Retrieved under: http://www.isoco.com/isococom/whitepapers/files/SemanticWeb-whitepaper-137.pdf

[BernLee01] Berners-Lee, T, Hendler, J., and Lassila, O.: *The Semantic Web*, Scientific American, May 2001

[BernLee03] Berners-Lee, T, Hendler, J.Scientific Publishing on the semantic web. In: Nature (http://www.nature.com/nature/debates/e-access/Articles/bernerslee.htm).

[CMUZZ96] Cole, R., Mariani, J., Uszkoreit, H., Zaenen, A. and Zue, V. (Eds.) *Survey of the State of the Art in Human Language Technology*, Chapter 8.4: ”Machine Aided Human Translation”. See http://www.cse.ogi.edu/CSLU/HLTsurvey/, February 1996.

[Declerk,Uszkoreit03] Declerck T., Uszkoreit H., “State of the art on multilinguality for ontologies, annotation services and user interfaces”, Deliverable Espeonto Project IST-2001-34373, January 2003

[vHa97] v. Hahn, W.: Providing Multilingual Term Explanations in Machine Aided Translation. In: Proceedings of the EAMT Workshop Kopenhagen 1997. S. 78 - 90.

[vHaAn94] v. Hahn, W. and Angelova, G. *Providing Factual Information in MAT*. In: Proc. of the Conf. "MT - 10 Years on", Cranfield, UK, November 1994, pp. 11/1 - 11/16.

[vHaJek00] v. Hahn, W. and Jekat, S.: *Multilingual VERBMOBIL-Dialogs: Experiments, Data Collection and Data Analysis*. In: W. Wahlster (ed.), Verbmobil: Foundations of Speech-to-Speech Translation. Berlin 2000. S. 575 – 582.

[vHaTes00] v. Hahn, W. and Tessiore, L.: *Functional Validation of a Machine Interpretation System: Verbmobil*. In: W. Wahlster (ed.), Verbmobil: Foundations of Speech-to-Speech Translation. Berlin 2000. S. 611 - 634 (mit Lorenzo Tessiore)

[HuSo92] Hutchins, John and Harold L. Somers: *An Introduction to Machine Translation*. London 1992

[JurafMartin00] Jurafski D., and Martin J., “An introduction to Natural Language Processing, computational Linguistics and Speech Recognition, Prentice Hall, New Jersey, 2000

[KBMT89] Center for Machine Translation. Carnegie Mellon University. *KBMT-89 Project Report*. Second Printing 1989

[KnightLuk94]] Knight K. And Luk S., “Building a Large-Scale Knowledge Base for Machine Translation”, Proceedings of the american association of Artificial Intelligence Conference, AAAI-94, WA, 1994

[NiCar92] Nirenburg, S., Carbonell, J., Tomita, M. and Goodman, K.: *Machine Translation: A Knowledge-Based Approach*. Morgan Kaufmann San Mateo1992.

[NoHaf97] Noy, N.F. and C.D. Hafner. *The State of the Art in Ontology Design: A Survey and Comparative Review*. AI Magazine, Vol. 18, No. 3, Fall 1997, pp. 53 - 74.

[Rus95] Russel, S. and Norvig, P.: *Artificial Intelligence. A Modern Approach*. Englewood Cliffs 1995.

[Sowa84] Sowa, J. *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, MA, 1984.

[Tru99] Trujillo, A.: *Translation Engines*. Springer London 1999

[VossenBloksma98] Vossen P., and Bloksma L., “Categories and Classifications in EuroWordnet”, Proceedings of LREC, 1998, Granada

[Wah00] Wahlster, W. (ed.): *Verbmobil:Foundations of Speech-to-Speech Translation*. Springer Berlin 2000.