IJCAI 87

PROCEEDINGS OF THE TENTH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE

MILAN • AUGUST 23-28, 1987

VOLUME 2

In Proceedings of the Tenth International Joint Conference on Artificial Intelligence (IJCAI87)

Direct Memory Access Translation[1]

Hideto Tomabechi[2]
Center for Machine Translation
Carnegie Mellon University
Pittsburgh, PA 15213 U.S.A.
412–268–6595
tomabechi@nl.cs.cmu.edu

Abstract

Direct Memory Access Translation (DMTRANS) is a theory of translation developed at CMT of CMU in which translation is viewd as an integrated part of cognitive processing. In this paradigm, understanding in source language is a recognition of input in terms of existing knowledge in memory and integration of the input into the memory. Context of sentences are established as what is left in memory after understanding previous sentences (or a preceding part of a sentence). Decisions made during translation are influenced by what is dynamically modified in memory through preceding recognitions. Since knowledge in memory is directly shared with the rest of cognition, during translation other cognitive processes such as inference can dynamically participate in the translation process.

I. Introduction

The Direct Memory Access Translation (DMTRANS) is a new approach to machine translation currently researched at the Center for Machine Translation (CMT) of CMU. We claim that every part of cognition dynamically participates in translation (as in any other cognitive process) through shared memory, and that a translation system aiming at fully-autonomous machine translation should be designed with this in mind. This project is an experimental project currently being developed at the CMT as a new generation MT system and should not be confused with the ongoing CMU-MT project (Tomita&Carbonell[1987]).

The current implementation of DMTRANS uses the spreading activation model as a simulated parallel memory search[3] to recognize input in terms of the existing knowledge in memory. Similar approaches to understanding languages are found in Quillian[1969], Collins[1969], Fahlman[1979], Riesbeck&Martin[1985]. Related past works in this area include Hirst[1982], Hahn[1983], Yokoyama&Hanakata[1986], and Charniak[1986]. We prefer this method, because translation is performed directly through the network of memory, which makes dynamic interaction with other memory-related processes possilble, and because all previously created memory structures can potentially participate in translation. DMTRANS extends and integrates theories of direct memory access understanding into translation with consideration of cross-cultural questions that accompany the attempt. We view translation as locating existing memory structures under the source language that the text is referring to and generating text that refers to these memory structures in the target language[4]. Often, a single memory structure is not shared by different languages and in that case, use of similar existing memory structures and explanation by surrounding memory structures replace direct generation from identified memory structures. Currently, the system is developed to translate between English and Japanese and is capable of understanding and generating fairly complex sentences between the two languages.

II. Where most MT systems fail

A. Ambiguation – No choice over others

Because most MT systems do not understand what they are translating, they are incapable of making decisions based on the content of the material they are translating. For example, the famous structurally ambiguous examples such as "I saw a man with a telescope" and "The man left by the door rotten" are handled by current systems by representing multiple interpretations of the input; however, this does not mean these systems are capable of handling garden path sentences, since none of these systems are capable of choosing the most correct interpretation over the others. Since inputs are translated sentence by sentence, virtually no contexual help is available during translation. This makes an autonomous translation extremely unlikely, because very often sentences can have multiple interpretations (most of which, humans are unaware of); without human assistance, such systems are incapable of selecting one interpretation over others[5]. Thus, being able to generate all possible interpretations of an input sentence does not automatically mean the system is capable of handling syntactically ambiguous sentences. We claim that the system should be able to select the correct interpretation (what speaker intended) in order to claim that it "handles" such a sentence. Unfortunately, most current MT systems fail in this task. By the same token, most MT systems fail in handling semantically ambiguous sentences. Consider the examples: "The quality of this paper is terrible" and "John gave Mary a punch". In the former example, the interpretation of paper should be different (for example, Japanese for 'thesis' and 'a sheet of paper' is different) according to what has been said before (or perhaps, visual perception of the situation may supply help). In the latter sentence, interpretation should be different again due to

[1] A significant part of this research was done at the Computer Science Department of Yale University as a multi-lingual natural language module of the PROJECT IVY (case-base medical consultation system supported by the National Library of Medicine under Grant No. 5-R01-LM04251).

[2] The author is supported by the Fulbright Scholarship under IIE Designation DGB, No. 15850622.

[3] A guided spreading activation is performed directly on the memory net and no modular syntactic analysis (Birnbaum[1986]) is done.

[4] Since understanding is done as accomodating input with already existing knowledge in memory (or past cases) we can also view DMTRANS as a kind of case-based translation theory.

[5] This problem is conspicuous when a sentence has a fairly complex structure including conjuncts. Consider "Show me the picture of lung with small cell carcinoma with magnification of ten and the brain with squamous cell carcinoma with magnification of five".

the context (Japanese for punch as PROPEL and punch as a drink is different). Again, being able to generate multiple interpretations of sentences does not mean the system is capable of handling semantically ambiguous sentences. The system instead should be able to choose appropriate interpretations.

B. Ellipses, Anaphora, Indirect Speech Act

In most MT systems, ellipsis in a sentence results in either no parse at all or output with missing slots. For example, in translating "kouryo suru to ittaga, totemo shinjigatai" ([he] said, [he] will consider [it], but [I] can hardly believe [it]) which is a typical Japanese sentence with missing subjects, most MT systems simply fail in filling in missing information[6]. Another example is "How often does squamous cell carcinoma metastasize to the brain? Lung? Large cell carcinoma?". Unless MT systems perform some strong inference at run-time, it is beyond their capacity to handle this phenomenon.

Anaphoric expressions are another kind of phenomenon that most MT systems fail to handle. Consider the example of "Musashi threw an arrow at the giant rat. It ate it." Current MT systems are satisfied with translating 'it' as 'it'[7]; however, this often creates problems: for example, Japanese does not prefer 'sore' (it) for animate objects whereas English refers to both animate and inanimate objects with 'it'. In some languages, the morphology of 'it' changes according to what it is referring to. In this sense, anaphora is another phenomenon most MT systems avoid.

Indirect speech acts also cause failure in most MT systems. At best, these systems output two possible interpretations of the utterance: the primary illocution and the secondary illocution, however, no preference for one over the other is made. A conference interpreter will take "Can you move over a little, your shoulder is blocking the picture" almost undoubtedly to be a request instead of a question. Without knowledge of what is it that the interpreter is translating, such an automatic choice is impossible.

III. What DMTRANS can do

DMTRANS outperforms most systems in choosing an appropriate interpretation of sentences over others in accordance with contexts. DMTRANS does not even realize many of the unlikely interpretations of the text (just like humans do not realize unlikely interpretations of an input text). This is possible because sentences are always recognized in context in DMTRANS, by performing strong predictions based on what has been recognized previously.

A. Contexual Recognition of Concepts

First, a brief view of the DMTRANS marker passing mechanism is in order. We have three kinds of markers[8] that are spread around in the memory network: the Activation-Marker (A-Marker), the Prediction-Marker (P-marker), and the Context-Marker (C-Marker). The A-Marker is to mark concepts (and the abstractions) that are being identified with inputs. The P-Marker is used to predict the next likely concept to be recognized, through knowledge of the possible sequence of concepts. The C-Marker is used to mark concepts that are likely to be input under a given context. When a word comes in, the word sends activation to (put an A-Marker on) a concept that the word is attached to and the activation is sent above the abstraction hierarchy in the network. The A-Marker contains the source of activation to indicate which concept originated the activation. The P-Marker contains the origin of prediction of a concept that the P-Marker is put on. Predictions are initially made (P-Markers are put) on all the first elements of concept sequences[9], and if a predicted concept receives activation (when A-Marker and P-Marker meet) then the next element of the concept sequence is predicted. A concept sequence is a sequence of concepts that represents an order of concepts that is unique to a language and is stored in root concepts[10]. When the last element of a concept sequence is activated, then the concept sequence is accepted and the associated root concept is recognized. When this happens, DMTRANS searches for (or creates if it does not exist yet) some concept underneath the root concept in the abstraction hierarchy that represents the specific input concept sequence[11] and activates that concept (another spreading activation). The C-Marker is stored in concepts (not necessarily root concepts) that influence the context of the text, and is sent to associated concepts when this concept is activated[12]. When activation is spread upward in the abstraction hierarchy and if more than one route exist (such as two meanings for a word), then the route through the C-Marked concepts are chosen unless the route hits a higher level concept that indicates a contrary preference.

In order to demonstrate this mechanism, let us examine a short translation of a semantically (word-sense) ambiguous sentence: "John is at IJCAI-87. He said the quality of the paper is terrible" (Figure 1). Initially, all the first elements of concept sequences (indicated by <...>) are predicted. The first word "John" comes in and activates the concept 'John' (put A-Marker on it) then the A-Marker is sent upward until it hits the concept 'person' which is predicted by 'at-person-loc' as the first element of the sequence. Then the prediction is sent to 'is' which gets activated by receiving A-Marker from next input word "is". Then 'at' is predicted as the third element of the sequence which meets activation from the input "at". Then the prediction for 'IJCAI-87' is made. When the word "IJCAI-87" comes in, and activates 'IJCAI-87' which was predicted as the last element of the concept sequence <person is at location>, this concept sequence is accepted and the root-concept 'at-person-loc' gets activated. Then the search is performed to find a specific concept under the root concept that indicates the input[13], and a concept refinement is conducted to get to 'at-John-IJCAI-87'. If this is not found, DMTRANS creates this concept as a specific episode of 'at-person-loc'. At the same time, since 'academic-conference' (activated by 'IJCAI-87') is a contexual-root concept it sends C-Markers to 'person-present-thesis', 'person-criticize-thesis', 'thesis', 'proceedings', etc.. When the next word "He" comes in, it sends activation upward and finds that the only male per-

---

[6]Simple heuristics such as "assume the missing subject to be the subject of the former clause' does not work here.

[7]As long as 'it' is translated as 'it' (perhaps 'sore' in Japanese), translation is treated as accurate in most systems.

[8]A-Marker and P-Marker are due to Riesbeck&Martin, which describes a more detailed picture of the way these two markers are passed around in memory.

[9]We use the term 'concept sequence' to represent some known sequence of concepts such as <feature, physical-object> which includes sequence of abstract concepts as in MOP components and also low level phrasal templates such as described by Becker[1975], Wilensky[1981], and Hovy[1986].

[10]Root concept is a concept that packages another concept in a structure, ie, MOP in DMTRANS. Verbs in a case-frame based lexicon are comparable structures.

[11]This is called 'concept refinement'. Lytinen discusses a rule-based version of this scheme.

[12]Concepts that receive the C-Marker include: Participants of a MOP, concepts representing events, explanation-patterns attached to a MOP.

[13]Concept refinement in DMTRANS is performed as a search for a node that packages the input recognized concept with links parallel to the links from the accepted root node to the elements of the accepted concept sequence.

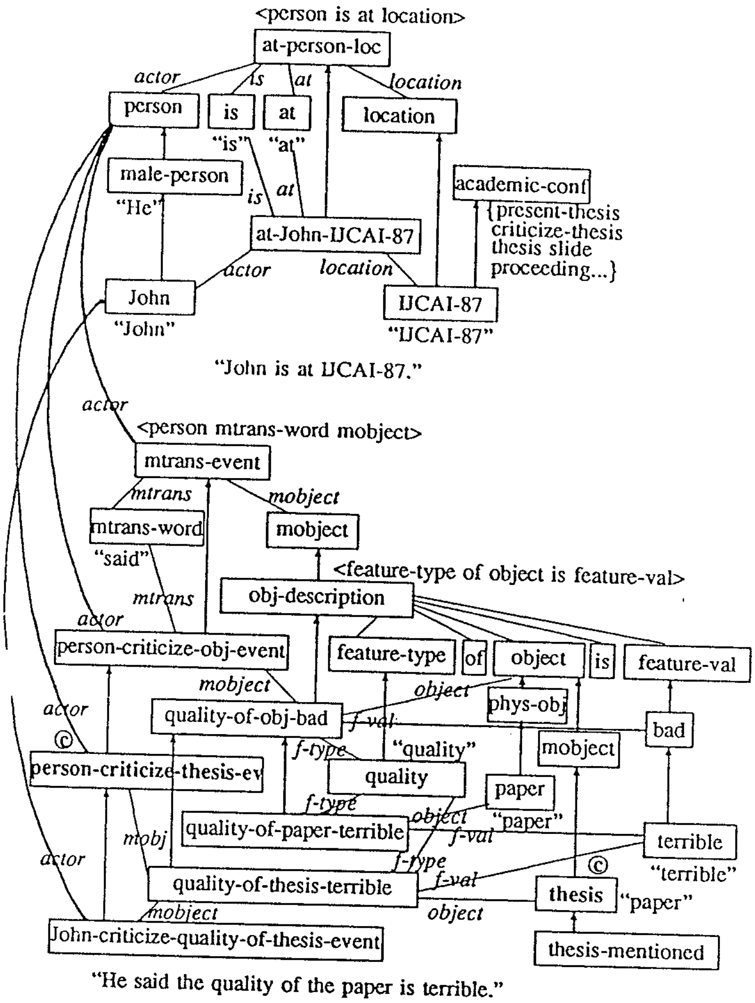

Figure 1: Concept refinement and context marking

son activated in memory is 'John', and activates 'John' again; 'person' gets re-activated, which is predicted as the first element of 'mtrans-event', then "said" comes in and fits as the second element of the concept sequence attached to 'mtrans-event'. Likewise, "The quality of this paper" is accepted, being identified with the sequence <feature-type of object is feature-value> attached to 'object-description'. One thing that happens ; that when "paper" which is attached both to 'paper' and 'thesis' comes in, only 'thesis' sends activation upward because 'thesis' was C-Marked by 'academic-conference' and 'paper' was not marked. This choice is not challenged when 'mtrans-event' is accepted and is concept-refined to 'person-criticize-thesis-event', since this concept also supports the contexual interpretation of "paper"[14]. This way, understanding is left as activated memory structures representing 'at-John-IJCAI-87' and 'John-criticize-quality-of-thesis-event' that are instances of the refined concepts under accepted root concepts.

Also, if two conflicting choices of a concept are marked by two C-Markers, the C-Marker put by the concept activated more recently gets preference. For example, in "John was writing a letter on a plane to IJCAI-87. The ink smeard. He said the quality of this paper is terrible" and in "John was printing a paper for IJCAI-87. The printer jammed. He said the quality of this paper is terrible", both 'paper' and 'thesis' are C-Marked by 'IJCAI-87'[15] and 'ink', 'IJCAI-87' and 'printer' respectively[16]. However, since, "ink" and "printer" both come after "IJCAI-87" in both cases, 'paper' is prefered over 'thesis' in both cases, and it gets activated. Unless these activations meet contradicting hypotheses elsewhere, 'paper' becomes the contextual interpretation of "paper".

B. Explanatory Generation

We have two different concept sequences stored in each root concept, one for English and one for Japanese. Especially because they represent texts from different language families, the sequences are rarely the same; however, the roles are shared, it is because memory structures are independent of languages and the types of roles are inherent in the root concepts, not in the languages. Similar approaches are taken in Lytinen[1984]'s MOPTRANS and CMU's current MT system. Both systems take advantage of shared memory structures for translation, the former using MOPs as the shared structure and the latter using case frames as the shared structure.

Generation begins with the result of memory activation parsing from input in one language. For each concept refined nodes left in memory, we do the following. 1) Check at the lexical node for the refined concept in the target language and if a lexical entry is found, generate in accordance with templates stored with the concept and we are done. 2) If not, which is the often the case[17], we generate according to the stored concept sequence for the target language. That is to generate from the first element of the concept sequences (go back to 1 with the first element of the concept sequence). 3) Since not all concepts have sequence attached to it, search the abstraction hierarchy upward for abstraction of the refined concept which has concept sequences attached to it. 4) Get the sequence from this abstraction and then instantiate with the roles in the refined-concepts. Then from the first element of the instantiated concepts sequence, repeat from looking up lexical node again. If not found, repeat from the 2 again to explain this concept.

One sample short translation is translating the Japanese sentence: "Gionshoja no kane no koe, shogyomujo no hibiki ari" which is translated to be: "The sound of the bell at Gionshoja has the tone of 'shogyomujo' (impermanence of all phenomena in the world)". Note that DMTRANS outputs 'shogyomujo' as 'shogyomujo', and adds the explanation of the word in parentheses. This is because an English lexical entry for the concept representing 'shogyomujo' was not found in memory. Since a concept may not be shared across languages, this type of explanation happens often, especially in a cross-cultural context[18]. If a lexical entry for the target language is not found, most MT systems simply halt excution. In contrast, DMTRANS outputs the explanation of the concept in the target language. This is possible through the explanatory generation mechanism described above. Explanation is performed by generation using the surrounding concepts with lexical entries in the target languages.

---

[14]C-Marked by the same contexual root concept as 'thesis'.

[15]Actually, C-Marked by 'academic-conference' which was activated by 'IJCAI-87'.

[16]These three concepts trigger (activate) contexual-root concepts.

[17]This is the inherent uniqueness of the DMTRANS system, that the system does not halt even if the lexical entry is not found in the target language: instead DMTRANS tries to explain the concept through surrounding concepts in the memory network that have lexical entries in the target language.

[18]The described explanatory generation mechanism works effectively in translation between English and Japanese, where a one to one match of concepts is often difficult to find due to the difference in the cultural contexts. Even words such as "river" and "kawa" (Japanese for river) which are normally substituted for one another without any further consideration, reveal difference in concepts attached to them, ie, the Japanese word "kawa" is normally associated with images of clear rapid streams. What about "kou" in Chinese?

C. Dynamic Interactions with the Rest of Cognition

Since translation is performed by directly accessing the memory network, other faculties of cognition can dynamically participate in translation. One example sentence here is "John threw an apple at the giant rat. It ate it". Whenever, a pronoun comes in as an input, DMTRANS tries to identify the object that is referred to[19]. In this example, the concept 'animal-ingest-object-event' gets activated by the input "it ate it". 'animal-ingest-object-event' is a MOP structure which is a kind of 'ingest-event'. It has two roles to be filled: Actor and Object. In order to determine the Actor, the inference mechanism is activated and it looks for activated concepts in memory that can be an Actor and finds 'giant-rat' to be a candidate given restrictions set forth by the MOP structure[20]. Then a search is made for concepts previously activated in memory that fit the requirements for Objects and 'apple' is selected to be an acceptable object of 'ingest-event'. This example only requires a minimum amount of work for deciding objects; however, this architecture allows for deeper inferences if necessary, such as utilizing causal relations stored in MOPs and eXplanation Patterns associated with higher level structures (Schank[1986])[21].

D. A Translation system that learns

DMTRANS is capable of creating new concepts while translating, and is capable of learning new vocabulary for newly created concepts in a multi-lingual context. When a concept refinement is performed, if a specific concept representing the input sentence is not found underneath the accepted root concept, a new specialization is created. Also, the user of the system is asked to input the English and Japanese names (words) for the concept (or input phrase can simply be stored as a phrasal lexicon). By the same token, we can simply assert facts to be translated by DMTRANS and the system stores the assertion as well as it translates it as long as it is not incompatible with what it already knows. At the same time, the acquired concept is accessible from different contexts because of the hierarchical organization of memory (Schank[1982]) that implements MOP structures. This way DMTRANS implements dynamic memory as its memory network and is capable of learning while translating.

IV. Conclusion

From a practical point of view, DMTRANS may be interesting because a lexically guided spreading activation mechanism is parallel in nature, and recent availability of massively parallel machines makes it an appealing theory for machine translation, utilizing such parallel architectures. However, the impact of this theory is that translation is performed as an integrated part of cognition, cooperating with other faculties through memory. Most MT systems have failed in tackling contexually ambiguous sentences; however, in DMTRANS, with use of episodic and thematic memory, and also the C-Marker passing, performance with ambiguous sentences is significantly improved. Explanatory generation handles culturally sensitive translations more effectively, especially when lexical entries in the target language are not available. Also, the dynamic participation of an inference mechanism contributes in handling phenomena such as anaphora, ellipsis, and indirect speech acts. A future possibility is that we may supplement DMTRANS with other input output channels to make the system's abilities closer to those of human interpreters in handling questions of pragmatics. In our understanding, memory is shared by all parts of cognition, and any cognitive task including translation should be dynamically assisted by every faculty with direct access to the memory.

[19] This is independent of the question whether to translate 'it' as 'it'. Even if we do, it is better to know what is referred by it with the reasons indicated before.

[20] If "John" is known to be a name of dog, we need more inference. Such as check the previously activated memory structure (propel-event) and infer where the apple is at now, etc..

[21] Actually, the understanding part of DMTRANS was originally designed as an integrated part of a case-based reasoning system to allow direct inference on input sentences.

Acknowledgments

Thanks are due to members of the PROJECT IVY, ie, Lawrence Hunter and Alain Rostain of Yale AI Project and Dr. Jerry Silbert of West Haven Veterans Administration Hospital. Dr. Christopher Riesbeck motivated me in the DMA paradigm. Also, many thanks to members of CMT especially to Eric Nyberg for valuable comments, and to Professor Jaim Carbonell and Dr. Masaru Tomita who are providing me an excellent opportunity to devote myself in machine translation research at the Center for Machine Translation.

References

[1] Becker, J.D. (1975) *The phrasal lexicon.* In 'Theoretical Issues in Natural Language Processing'.

[2] Birnbaum, L. A. (1986) *Integrated Theory of Understanding and Planning*, Ph.D. thesis Yale University.

[3] Charniak, E (1986) *A neat theory of marker passing.* In 'Proceedings of the AAAI-86'.

[4] Collins, A. and Quillian M.R. (1969) *Retrival time from semantic memory.* Journal of Verbal Learning and Verbal Behavior.

[5] Fahlman, S.E. (1983) *NETL: A system for representing and using real-world knowledge.* The MIT Press.

[6] Hahn, U and Reimer U. (1983) *World expert parsing: An approach to text parsing with a distributed lexical grammar.* Technical Report, Universitat Konstanz, West Germany.

[7] Hirst, G and Charniak, E. (1982) *Word sense and case slot disambiguation.* In 'Proceedings of the AAAI-82'.

[8] Hovy, Eduard (1986) Ph.D. thesis Yale University. Forthcoming.

[9] Lytinen Steve. (1984) *The organization of knowledge in a multi-lingual, integrated parser.* Ph.D. thesis Yale University.

[10] Quillian, M.R. (1969) *The teachable language comprehender.* BBN Scientific Report 10.

[11] Riesbeck, C. and Martin, C. (1985) *Direct Memory Access Parsing.* Yale University Report 354.

[12] Schank, Roger C. (1982) *Dynamic Memory.* Cambridge University Press.

[13] Schank, Roger C. (1986) *Explanation Patterns.* Lawrence Erlbaum Associates.

[14] Tomita, M. and Carbonell. J. (1987) *The Universal Parser Architecture for knowledge-Based Machine Translation.* In 'Proceedings of the IJCAI-87'.

[15] Wilensky, R. (1981) *Knowledge-based approach to language processing.* In 'Proceedings of the IJCAI-81'.

[16] Yokoyama, S. and Hanakata, K. (1986) *Conceptual Lexicon Using an Object-Oriented Language.* In 'Proceedings of the COLING-86'.