THE NEW GENERATION BEHAVIORTRAN : DESIGN PHILOSOPHY AND SYSTEM ARCHITECTURE

*Yu-Ling Una Hsu and *+Keh-Yih Su

*Behavior Design Corporation
No. 5, 2F, Industrial East Road IV
Science-Based Industrial Park
Hsinchu, Taiwan 30077, R.O.C.

+Department of Electrical Engineering
National Tsing-Hua University
Hsinchu, Taiwan 30043, R.O.C.

Email: *una@bdc.com.tw, *+kysu@bdc.com.tw

ABSTRACT

In many conventional machine translation systems, the translation outputs are usually strongly affected by the syntactic information of the source sentences and thus tend to produce literal translations that are not natural to the native speakers. In this paper, we introduce the design philosophy and system architecture of the new generation BehaviorTran, which will enable an MT system to operate with high modularity and to acquire its translation knowledge from a bilingual corpus with a two-way training method. In such a paradigm, the knowledge bases only provide static descriptions on the legal forms of the constructs, while ambiguity resolution and preference evaluation are governed by sets of statistical parameters. This makes it easier to adapt the system to specific user styles and maintain different parameter sets for different customers. Thus, it is expected to be a promising paradigm for producing satisfactory translations.

1. Introduction to the First Generation BehaviorTran

The BehaviorTran English-Chinese Machine Translation System (formerally ArchTran) is the first of its kind research launched in Taiwan, and is among the few commercialized English-Chinese systems in the world (Chen[1], Wu[2]).

The research on BehaviorTran began as a joint project between National Tsing Hua University, Taiwan, and Behavior Tech Computer Corporation (BTC) in May, 1985. And as the scope of research gradually extended, the BehaviorTran was later transferred to Behavior Design Corporation (BDC) in Feb., 1988 to continue the improvements on the system.

After four years of research, the first generation BehaviorTran was released in 1989 and serves as the kernel of a value-added network (VAN)-based translation service. The system is running on the SUN workstation and written in C language. Its primary domain is computer manuals and related documents.

The overall translation strategy adopted in the first generation BehaviorTran is conventional transfer-based approach. In this approach, the whole translation process can be logically divided into three phases, namely, analysis, transfer, and synthesis. In the first generation BehaviorTran, the English analysis component consists of a set of ATN-style augmented context-free phrase structure rules, which will parse the input English sentences into corresponding syntactic trees (Hsu [3]). And the transfer and synthesis operations are encoded in a set of pattern-action pairs, called tranfer rules, to carry out a sequence of tree to tree mappings to reflect the changes in substructures and linear order in the source-target language pair (Chang [4]).

2. Motivation for the Revision of BehaviorTran

Since the release of the first generation BehaviorTran, BDC translation center has established a customer base of several internationally-renowned companies. From several years of practical experience and the feedbacks from posteditors and customers, we find some drawbacks of the original system which urge us to make a thorough revision to the first generation BehaviorTran. The major drawbacks are stated as follows:

First, the degree of modularity in the first generation BehaviorTran is low. The application of transfer rules and the selection of target translations are closely related to the output of source language analysis grammar. Thus, once the analysis grammar is modified, a great number of transfer rules or lexical information should be modified accordingly. That greatly increases the load of system maintenance. Moreover, since different components are intricately related, when a new source or target language comes into play, most parts of the original components cannot be reused. This drawback becomes more and more sailent since BehaviorTran intends to extend itself to a multilingual translation system.

Second, as mentioned previously, the transfer rules in the first generation of BehaviorTran are mainly based on the output of superficial syntactic parse trees of the analysis grammar. However, since parse trees are usually huge and branchy, and sentences similar in meaning may be presented in different surface syntactic structures, the transfer operations required for producing good translations are numerous and usually very complex and complicated. Thus, it is hard and costly to acquire a complete set of transfer rules and to ensure correct interactions among them.

Third, since the transfer operations in the first generation BehaviorTran are very complex, system designers usually tend to use minimal numbers of local adjustments in the transfer phase to get readable target translations. As a result, the output target structures usually retain a large portion of the source information, such as the part of speech of terminal words and the sentence patterns. The transfer mapping, thus, may minimize the required transfer operations, but may not optimize the translation quality. Consequently, literal translations which are not natural enough to native speakers are generated from time to time

Fourth, except the lexical tagger, which uses statistical information to solve lexical category ambiguities, most knowledge bases of the first generation BehaviorTran are written by linguists.

However, as the system scales up, this kind of rule-based approach suffers from many problems as indicated below:

- It is hard to maintain consistency of the large amount of fine-grained knowledge among different persons at different time.
- It is hard and costly to acquire the large amount of fine-grained knowledge with human intervention.
- It is hard for human to deal with complex and irregular knowledge in terms of formal and precise rules. Exceptions of rules occur from time to time.
- It is hard to maintain uncertainty knowledge due to the lack of objective preference measure.

Fifth, as the business of the BDC translation center grows, the translation domains of BehaviorTran extend from computer science to electrical engineering, mechanical engineering, aviation, navigation etc. and the number of customers and posteditors are increasing. It becomes more and more sailent that the special patterns and style in each subdomain should be taken into consideration to render satisfactory translations. Besides, the feedback from posteditors and customers should also be incorporated to improve the translation system. However, in the first generation BehaviorTran, the work of sublanguage knowledge acquisition and feedback analysis is labor-intensive and, thus, very time-consuming and not cost-effective.

Since fixing the drawbacks mentioned above requires a revolutionary change of the design philosophy and basic architecture of the first generation, we started to develop the new generation BehaviorTran.

3. Design Philosophy of the New Generation BehaviorTran

3.1 A Cooperative Approach Integrating Both Linguistic and Statistical Information

To avoid the shortcomings of rule-based approach, the design philosophy of the new generation BehaviorTran moves toward a corpus-based, statistical-oriented approach. With this approach, linguists are requested to construct the language model, corpus are used as the main information source and statistical techniques are used to learn model parameters and automatically acquire the knowledge from the corpus. The advantages of this approach are listed below:

- uncertainty or preference is interpreted objectively and consistently
- consistency can be easily maintained even in large scale systems
- automatic training is possible with least human intervention
- well-established statistical theories and techniques are available
- remove the burden of rule induction from linguists to machine
- easy to meet the desirable designing goals of wide coverage, robustness, adaptability, controllability, parameterization and cost-effectiveness

Our researches along this line of design philosophy include a bi-directional transfer model (Chang [5]), various kinds of score functions for selecting the best candidate (Su [6], Chang [7], Chiang [8], Lin [9]), semi-automatic grammar construction (Su [10]), compound extraction (Wu [11]), etc.

3.2 Introduction of Intermediate Normal Forms

Another major change in the new generation BehaviorTran is the introduction of the Normal Form (NF) levels. NFs refer to the intermediate structures between source language parse trees and target language output translations. With the introduction of NFs, we intend to set up a set of linguistically-justified intermediate levels which can separate the original transfer process into several independent phases, and can serve in a manner relatively independent of involved language paris and surface forms. This idea is close to traditional interlingua approach. However, our NFs are unlike interlingua since they are not universal representations of all languages. Instead, NFs are normalized language-specific representations minus the language-specific idiosyncracies, which are most troublesome in MT. NFs also contain the (near) universal representation of semantic roles and relations. In this sense, NFs are similar in spirit ot the reduced f-structre in Lexical Functional Gramar (LFG), which can be directly mapped to a (potential universal) semantic representation (Halvorsen [12]). Besides, another major difference between NFs and interlingua is that NFs do not involve the decomposition of lexical entries into semantic primitives (e.g. "kill" = "cause to become not alive" (Schank [13])). It has been pointed out in many MT systems (Bennett [14], Durand [15]) that the set of universal semantic primitives are hard to be clearly defined. And it is not obvious to us that the decomposition of words into primitives will improve the quality of translation.

A schematic view of the translation flow in the new generation BehaviorTran is shown in Figure 1 below. PT stands for the parse tree, NF1 stands for the first-level normal form, and NF2 stands for the second-level normal form. The subscripts 's' and 't' stand for the source and target language respectively. P(X|Y) represents the conditional probability for X to appear given that Y is observed. Such parameters (conditional probabilities) are used to assign preference scores for disambiguation.

We further assume that the parse trees are produced based on a phrase structure grammar G, the NF1 constructs are produced based on a set of normalization rules, NR1, and the NF2's are produced according to a second set of normalization rules, NR2. In addition, the reverse operations are directed by sets of generation rules of the various levels (GR2, GR1, and GR0), which specify the sets of legal $NF1_t$, PT_t and T's in the generation processes.

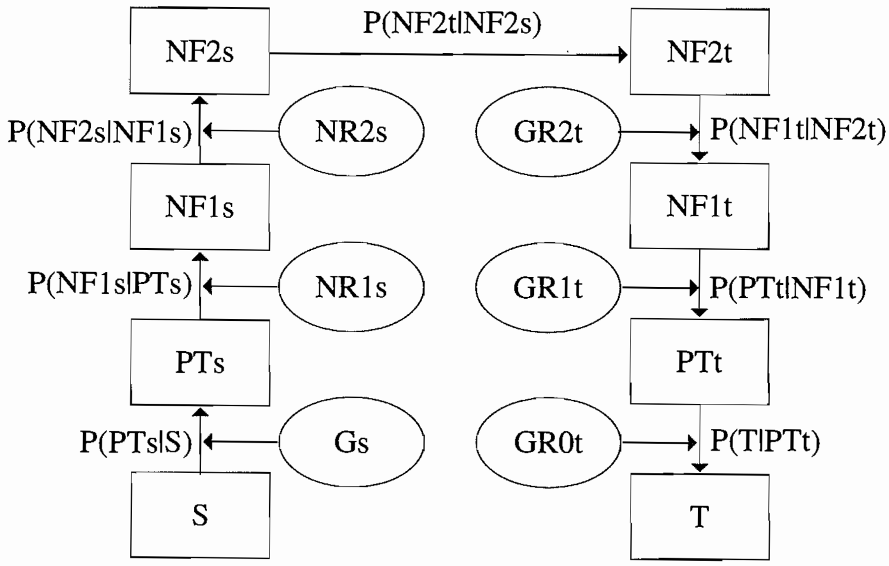

Figure 1 A Schematic View of the Translation Flow in BehaviorTran

Note that we introduce two intermediate structures in each language. NF1 is the level for syntactic normalization, in which all the elements that do not influence the cognitive meaning of a sentence will be eliminated. Thus the function words, punctuations, and unnecessary branchings and nodes are eliminated in NF1. NF2 is a semantically-oriented representation in which the basic constituents (i.e. governers, dependants, and modifiers) are marked with their semantic case roles (e.g. Agent, Theme, Time, Manner etc. (湯 [16], 詞庫 [17])), and some closed class elements (e.g. tense, aspect, modality, case markers, etc.) are extracted and recorded as a set of attribute-value pairs on the relevant nodes. Details about the NFs and their merits are illustrated in section 4.

3.3 Two-Way Training

The goal of a practical MT system is to produce fluent outputs that are natural to the native speakers of the target language. However, under the traditional transfer-based MT architecture, most output translations are strongly influenced by the sentence patterns of the source language and many literal translations are produced across the transfer phase (Somers [18], Su [19]). Such source-dependency is easily introduced to a transfer-based MT system in the one-way analysis, transfer and generation flow as mentioned earlier.

An alternative approach we propose is a two-way training approach which acquire the translation knowledge from a bilingual corpus. The bilingual corpus contain lots of well-polished source-target sentence pairs which are, undoubtedly, wonderful sources for transfer knowledge acquisition.

To change the system architecture from one-way design toward two-way design, the transfer knowledge should thus be trained from both properly normalized source and target knowledge representations, which should both fall within the range of the sentences that will be produced by the native post-editors, according to the discourse context of the source language and the target language respectively. The following flow shows the general idea for training a two-way system. The bold arrows at the right hand side emphasize that the intermediate representations for the target language are directly derived from the target sentences in an aligned bilingual corpus.

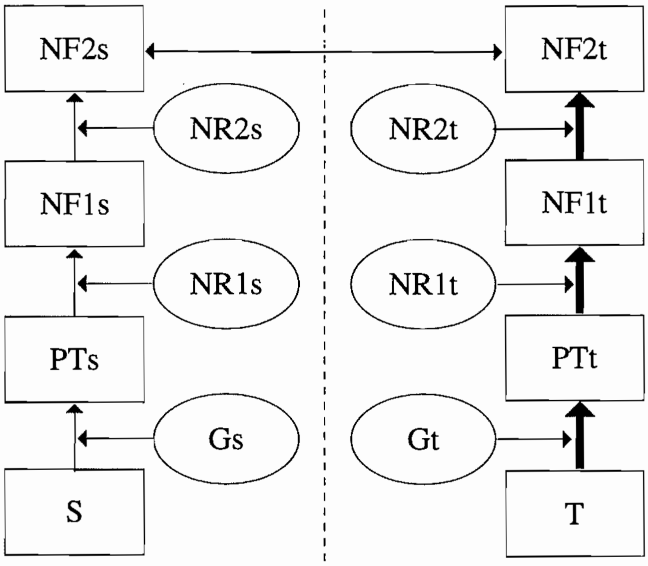

Figure 2 Two-Way Training Flow

Note that, the translation flow still follows the analysis, transfer and generation steps, but the training procedure for knowledge acquisition is different from the one-way design system. The arrow symbols indicate that the PT's, NF1's and NF2's for *both* the source and target sentences are derived from the source and target sentences respectively, based on their own phrase structure grammars and normalization rules. Thus, all such intermediate representations are guaranteed to fall within the range of the sentences that will be produced by the native speakers of the source and target languages; the transfer phase only *select* those preferred candidates among such constructs. In addition, the transfer parameters are estimated based on such intermediate representations and the transfer knowledge is derived from both the source and target sentences of an aligned bilingual corpus. Details about the formulations and training issues of the two-way training model can be found in Su [20].

4. System Architecture of the New Generation BehaviorTran

4.1 New System Architecture

With all the new ideas mentioned above, the new generation BehaviorTran gradually takes shape. The schematic veiw of the new architecture has been shown in Figure 1. Since the current interest of BehaviorTran is the language pair English-Chinese, detailed architectures of the English-to-Chinese and Chinese-to-English translation flow are presented in Figure 3 and Figure 4 respectively. The English-Chinese translation flow is briefly illustrated with an example in section 4.2 (許 [21]).

Figure 3 English-Chinese Machine Translation Flow

ENF2 → CNF2
1. structure selection
2. lexical sense selection

ENF1 → ENF2 (English analysis)
1. extract singular/pural information of countable nouns
2. extract definite/indefinite information of NPs
3. identify correct sense of each word
4. identify semantic cases of arguments and modifiers
5. extract case markers
6. extract tense, aspect information
7. eliminate voice information
8. insert and coindex omitted arguments
9. extract sentence type information
10. extract conjunctions

CNF2 → CNF1
1. select appropriate expression for each sense
2. determine the surface forms of lexical NP, pronoun and pro
3. identify the anophors of pronouns
4. realize the definite/indefinite expression of NPs
5. realize number information of countable nouns
6. determine the realization of sentence type
7. determine voice of sentences
8. determine the realization of tense, aspect information
9. determine the realiztion of conjunction
10. determine the realization of case markers
11. determine syntactic realization of each constituents
12. determine linear order of constituents

EPT → ENF1
1. eliminate punctuations
2. eliminate function words
3. structure flattening
4. normalize syntactic labels

CNF1 → CPT
1. insert punctuations
2. insert function words
3. replace normalized syntactic labels with corresponding parse tree nodes
(4. structure expanding)

ES → EPT
1. identify split idiom
2. morphological analysis
3. parsing

CPT → CS
1. locate split idiom at right position
2. realize morphological information

Figure 4 Chinese-English Machine Translation Flow

ENF2 ← CNF2

1. structure selection
2. lexical sense selection

CNF2 → ENF2 → ENF1

1. select appropriate expression for each sense
2. determine the surface forms of lexical NP, pronoun and pro
3. identify the anophors of pronouns
4. realize the definite/indefinite expression of NPs
5. realize number information of countable nouns
6. determine the realization of sentence type
7. determine voice of sentences
8. determine the realization of tense, aspect information
9. determine the realiztion of conjunction
10. determine the realization of case markers
11. determine syntactic realization of each constituents
12. determine linear order of constituents

CNF1 → CNF2

1. extract singular/pural information of countable nouns
2. extract definite/indefinite information of NPs
3. identify correct sense of each word
4. identify semantic cases of arguments and modifiers
5. extract case markers
6. extract tense, aspect information
7. eliminate voice information
8. insert and coindex omitted arguments
9. extract sentence type information
10. extract conjunctions

ENF1 → EPT

1. insert punctuations
2. insert function words
3. replace normalized syntactic labels with corresponding parse tree nodes

(4. structure expanding)

CPT → CNF1

1. eliminate punctuations
2. eliminate function words
3. structure flattening
4. normalize syntactic labels

EPT → ES

1. locate split idiom at right position
2. realize morphological information

CS → CPT

1. word segmentaion
2. identify split idiom s
3. morphological analysi
4. parsing

4.2 English-Chinese Translation Flow

A. Syntactic Parsing

The first module is a syntactic parsing module which produces the parse tree (PT) of the input sentence according to the phrase structure grammar of the source language. This module provides the syntactic information for the input sentences. An example of the parse tree is shown in the following figure. ("To meet spectrum analyzer specifications, allow a 30 minutes warm-up before attempting to make any calibrated measurements .")

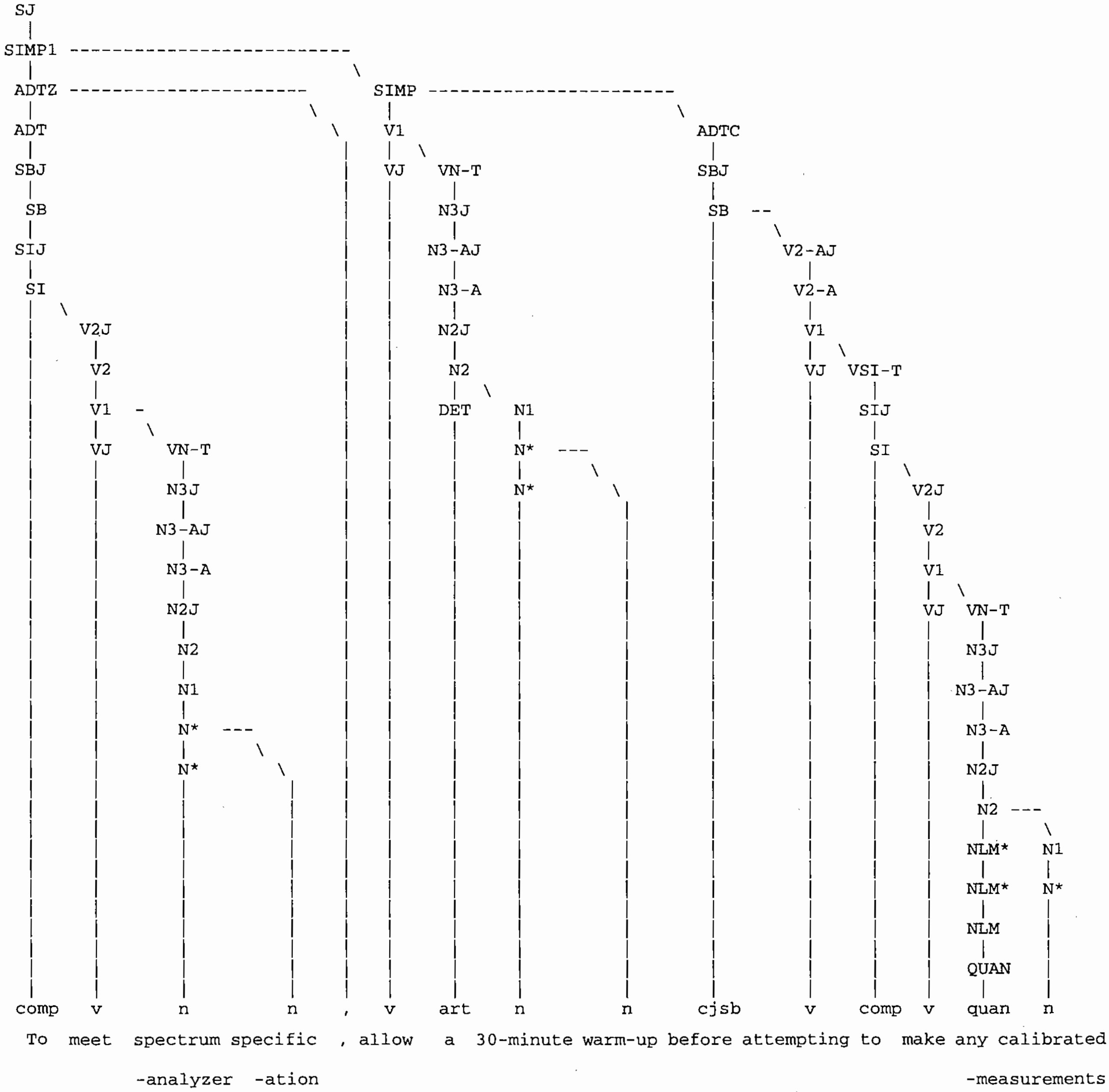

Figure 5 Example : A Source Parse Tree

Note that the parse trees produced by the phrase structure grammar of a large-scaled system are usually huge, branchy, and nodose. So it will be an arduous work to build the transfer grammar directly from the parse tree constructs.

B. Level 1 Normalization(NF1)

In NF1 level, all the elements that do not contribute to the cognitive meaning of a sentence are eliminated. Those elements, such as punctuations, function words, and unbranching tree nodes, will not influence the choice of target translations and shall not be taken into consideration in subsequent stages. Besides, removing those redundant information will reduce the size of the possible parameter space and will simplify the process of further normalization.

```
SIMP ------------------------------------------------------------
 |                                     \    \                         \
 SB -----------                        V    NP ---                    SB ---------------
 |    \        \                       |    |    \                    |        \        \
 |    V        NP ------               |    DET  NP -------           |        V        VP ----------
 |    |        |       \               |    |    |        \           |        |        |   \        \
 |    |        |        |              |    |    |         |          |        |        |   NLM      NP
 |    |        |        |              |    |    |         |          |        |        |   |        |
comp  v        n        n              v    art  n         n         cjsb      v        v   quan     n

 To  meet  spectrum  specific  allow   a 30-minute  warm-up  before attempting  make  any calibrated
           -analyzer  -ations                                                        -measurements
```

Figure 6 Example : A Source NF1 Tree

In the current example, the syntax tree is greatly compacted by retaining only the major syntax structure; a large number of nodes are compacted and re-labelled with representative node labels.

C. Level 2 Normalization (NF2)

The NF2 level is the level for semantic representation. A NF2 tree is an order-free dependency tree which specifies the semantic case roles of its governers (head), dependants (arguments), and modifiers (adjuncts), and is enriched with sets of feature-value pairs (such as tense, modality, voice, number etc.) on superier nodes.

The reason we perform the semantic-oriented normalization in NF2 is two-fold. First, as most MT researchers agree, what should be preserved in the process of translation is the semantic meaning of the source sentence instead of its syntactic structure. However, as mentioned earlier, in most traditional transfer-based MT systems, the transfer rules are constructed mainly based on the source syntactic trees, and therefore the translation outputs are usually strongly affected by the source sentence patterns and are often judged by the native speakers as "readable but not nature enough". Thus, elevating the intermediate representation from a syntactic parse tree to a semantic dependancy tree will make it possible to some extent to get rid of the tie from the syntactic information of the source sentence, and make it easier to render correct, fluent, and natural target translations.

Second, since NF2 involves feature extraction (i.e. remove some surface elements (e.g. modals, case markers, etc.) and record them as a set of feature-value pairs on superier nodes), some sentences that are different in their syntactic forms may be normalized to the same NF2 construct (e.g. active-passive pairs), and thus may further reduce the possible parameter space for statistical training. Since the parameters required to characterize the translation model may be numerous, the compression and normalization of the intermediate constructs is a very important processes which actually makes the two-way training approach feasible.

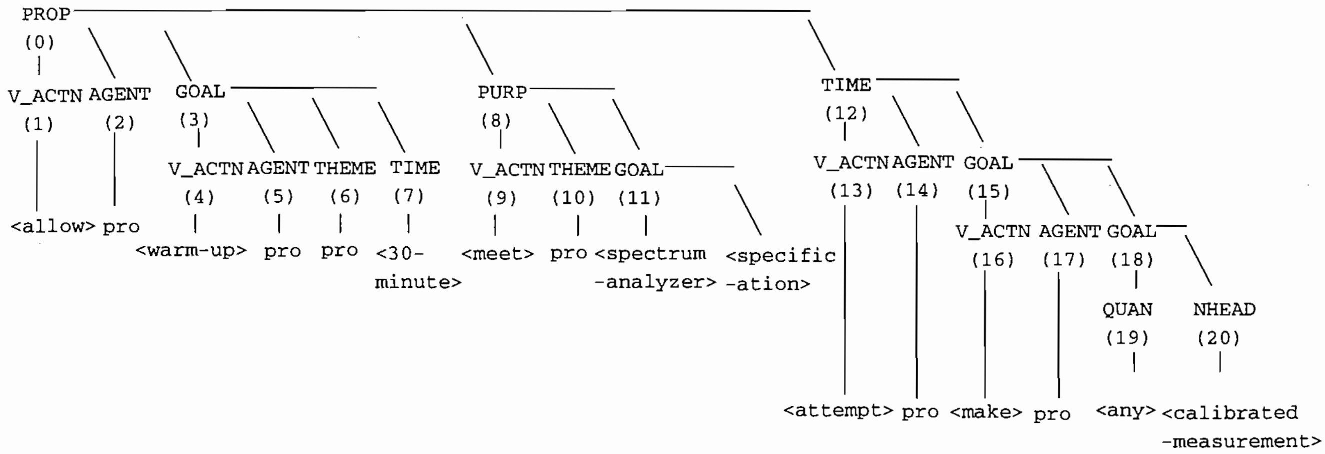

Figure 7 Example : A Source NF2 Tree

The example above is analyzed as a NF2 tree which specifies the Action (V_ACTN) being performed, the Agent (AGENT) who conducted the action, as well as the TIME, GOAL and PURPose for conducting the action (Extracted features are not presented here for simplification).

D. Target NF2 Selection

Given the NF2 tree of the source sentence, a proper NF2 tree of the target sentence could be selected among the set of target NF trees that are produced by the target analysis grammar. The selection could be made based on the parameters trained by the two-way training method, and can further incorporate the discourse and stylistic information. Note that the process is actually a 'selection' process rather than a 'derivation' or 'transfer' process from the source NF2 trees. By selection, the target NF2 is only selected from legal target NF2's, and therefore the output target NF2 will not be an illegal one.

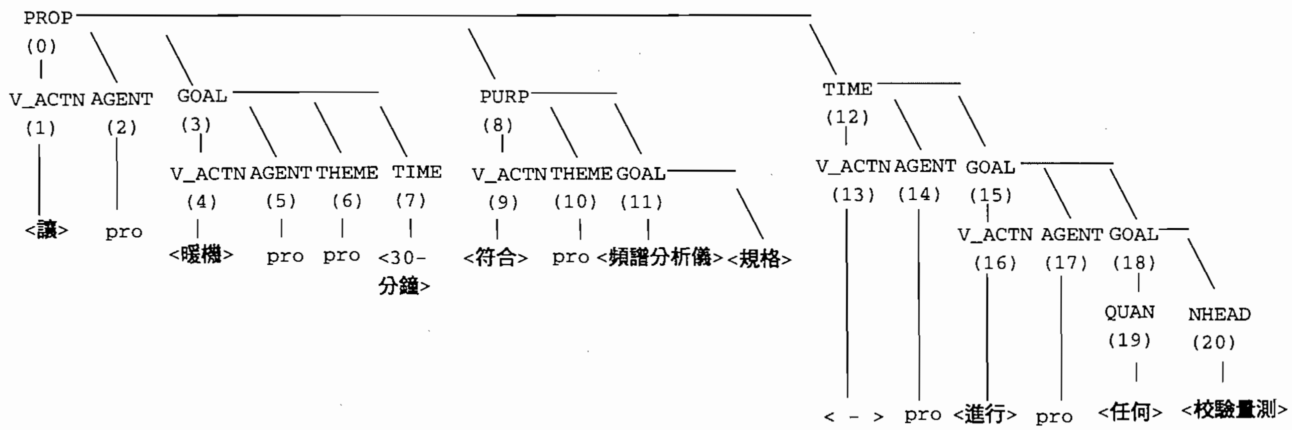

Figure 8 Example : A Selected Target NF2 Tree

In the above example, the selected target NF2 does not differ much from the source NF2 due to previous normalization. The major change here is the transfer of word senses (where a sense is represented with a pair of angle brackets).

E. Normalized Structure Generation

Given a target NF2 that is derived from the target grammar, the next step would be to choose an appropriate normalized syntax tree for generation. As in the phase above, we may also include discourse and stylistic information in this phase to select the realization form of feature-value pairs (e.g. case markers, tense, voice, etc.) and to select the preferred linear order of constituents. An NF1 tree of the target sentence generated in this way will contain the skeletal syntactic structure for the target sentence, as shown below:

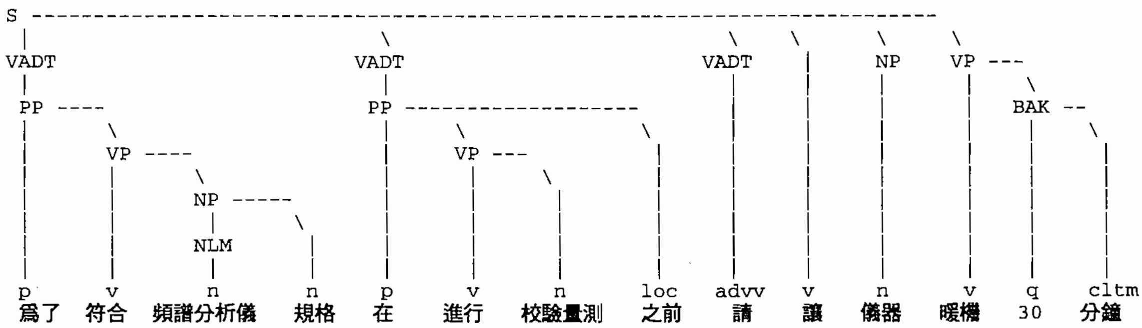

Figure 9 Example : A Selected Target NF1 Tree

Note again that the generation step here is actually accomplished by a selection process from a set of legal normalized syntax trees which are derived from the target grammar. Thus, the final translation output will not be deformed and produce unnatural translation.

F. Surface Structure Generation

After the NF1 tree is generated, the subsequent step is to determine whether some function words or punctuations should be added to improve its fluency and meet user-preferred style. The final syntactic tree of the example is shown below. (In this simple case, only two punctuation marks are patched here.)

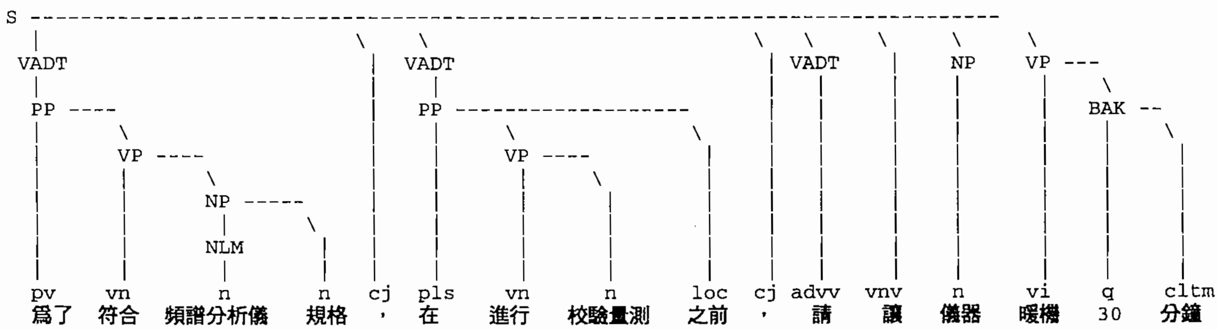

Figure 10 Example : A Selected Target Syntax Tree

G. Morphological Generation

Finally, the morphological generation is performed in the final step to generate the morphemes required in the target language. In the above example, no specific tokens of this kind are inserted, and the final preferred translation is "爲了符合頻譜分析儀規格，在進行校驗量測之前，請讓儀器暖機30分鐘。" The Chinese-specific morpheme 「們」, which is used in conjunction with certain nouns to produce their plural form, such as 「同學們」, for example, is generated in this phase.

4.3 Merits of the New Architecture

- ❑ With the introduction of NF representations, the output of the analysis grammar for any source language can be used to synthesize any number of target languages without rewriting the analysis component, and vice versa the generation component. Thus, new language pairs may be added to the MT system with a minimum amount of development time.
- ❑ NFs separate the transfer process into several phases. Operations in each phase are independent to those in other phases. This greatly enhances the modularity of the system and lighten the burden of manipulation and maintenance in each phase.
- ❑ The target normal forms are directly derived from the target grammar, not a deformed version from the source grammar, and thus can eliminate the bias resulted from the source language.
- ❑ The mapping between the source and target normal forms can be easily tuned by the two-way training method to generate the sentences which reflect the preferred sentence patterns and styles encoded in the training corpus.
- ❑ The knowledge bases in this architecture only provide static descriptions on the legal forms of the constructs, while ambiguity resolution or preference evaluation is governed by sets of parameters. This makes it easier to adapt the system to specific user styles and maintain different parameter sets for different customers.

5. Concluding Remarks

In this paper, we present the design philosophy and architecture in the new generation BehaviorTran. With its superiority in knowledge acquisition, modularity, adaptability, and bidirectionality, this new architecture is expected to play an important role in designing the MT systems of the next generation. And all these new changes enable BehaviorTran to gain more flexibility and better performance and to move from a purely English-Chinese translation system toward a multilingual translation system.

References

[1] Chen, S.-C., J.-S. Chang, J.-N. Wang and K.-Y. Su, "ArchTran: A Corpus-Based Statistics-Oriented English-Chinese Machine Translation System," *Proceedings of Machine Translation Summit III*, pp. 33--40, Washington, D.C., USA, 1991.

[2] Wu, M.-W., J.-S. Chang and K.-Y. Su, "The Current Status of ArchTran: A Corpus-Based Statistics-Oriented English-Chinese Machine Translation System," *Proceedings of the 1991*

Workshop on Machine Translation, pp. 123-138, Nantou, Taiwan, June 24-26, 1991.

[3] Hsu, H.-H. and K.-Y. Su, "A Bottom-Up Parser in the Machine Translation System with the Essence of ATN", *Proceedings of International Computer Symposium (ICS)* 1986, Vol. 1 of 3, pp. 166-173, Tainan, Taiwan, R.O.C., Dec 17-19, 1986.

[4] Chang, C.-L. and K.-Y. Su, "A New Mechanism of Transfer for An English-Chinese MT System," Conf. on Translation Today, the Hong Kong Institute for Promotion of Chinese Culture, H.K., Dec. 17-21, 1987.

[5] Chang, J.-S. and K.-Y. Su, "A Corpus-Based Statistics-Oriented Transfer and Generation Model for Machine Translation," *Proceedings of TMI-93*, pp. 3--14, 5th Int. Conf. on Theoretical and Methodological Issues in Machine Translation, Kyoto, Japan, July 14--16, 1993.

[6] Su, K.-Y. and J.-S. Chang, "Semantic and Syntactic Aspects of Score Function," *Proc. of COLING-88*, vol. 2, pp. 642--644, 12th Int. Conf. on Computational Linguistics, Budapest, Hungary, 1988.

[7] Chang, J.-S., Y.-F. Luo and K.-Y. Su, "GPSM: A Generalized Probabilistic Semantic Model for Ambiguity Resolution," *Proceedings of ACL-92*, pp. 177--184, 30th Annual Meeting of the Association for Computational Linguistics, University of Delaware, Newark, DE, USA, 1992.

[8] Chiang, T.-H., Y.-C. Lin and K.-Y. Su, "Robust Learning, Smoothing and Parameter Tying on Syntactic Ambiguity Resolution," to appear in *Journal of ACL*, 1995.

[9] Lin, Yi-Chung, Tung-Hui Chiang and Keh-Yih Su, "Discrimination Oriented Probabilistic Tagging," *Proceedings of ROCLING-V*, ROC Computational Linguistics Conference V, pp. 87--96, 1992.

[10] Su, K.-Y., Y.-L. Hsu and C. Saillard, "Constructing A Phrase Structure Grammar By Incorporating Linguistic Knowledge And Statistical LOG-Likelihood Ratio," *Proceedings of ROCLING-IV*, pp. 257--275, National Chiao-Tung Univ., Taiwan, August 18--20, 1991.

[11] Wu, Ming-Wen and Keh-Yih Su, "Corpus-based Automatic Compound Extraction with Mutual Information and Relative Frequency Count," *Proceeding of ROCLING VI*, pp. 207-216, Nantou, R.O.C., September 2-4, 1993.

[12] Halvorsen, Per-Kristian, "Semantics for Lexical-Functional Grammar," *Linguistic Inquiry* 14.4, pp. 567-615, 1983.

[13] Schank, R.C., "Identification of Conceptualizations Underlying Natural Language", in R.Schank & K. Colby (eds.) *Computer Models of Thought and Language*,, San Francisco:Freeman, pp. 187-248.

[14] Bennett, W.S., "METAL: Past, Present, and Future," *Proceedings of the 1990 Workshops on Machine Translation*, pp. 1-35, Nantou, R.O.C, June 24-26, 1991.

[15] Durand, J et.al. "The Eurotra Linguistic Specifications: An Overview." Machine Translation 6 (1991), vol. 6, no. 2, pp. 103 - 147.

[16] 湯廷池，「原則參數語法，對比分析與機器翻譯」，*漢語詞法句法四集*，學生書局，

頁 251-335。

[17] 中文詞知識庫小組，「中文詞類分析」，技術報告 (no. 93-05)，中研院資訊科學研究所。

[18] Somers, Harold, "Current Research in Machine Translation", *Machine Translation*, vol. 7, pp.231-247, Kluwer Academic Publishers, 1993.

[19] Su, K.-Y. and J.-S. Chang, "Why MT Systems Are Still Not Widely Used ? " *Machine Translation*, vol. 7, no. 4, pp. 285--291, Kluwer Academic Publishers, 1993.

[20] Su, K.-Y., J.-S. Chang and Y.-L. Una Hsu, "A Corpus-based Two-Way Design for Parameterized MT Systems: Rationale, Architecture and Training Issues," to appear in *Proceedings of TMI-95*, 1995.

[21] 許玉玲、郭明緯，「MT 新架構」，自然語言處理部門技術報告 (no. LG-94-01-NF-01)，致遠科技。