Automatic Addition of Verbal Semantic Attributes to a Japanese-to-English Valency Transfer Dictionary

Hiromi Nakaiwa
NTT Communication Science Laboratories
2-4 Hikari-dai, Seika-cho,
Soraku-gun, Kyoto,
619-0237, JAPAN
nakaiwa@cslab.kecl.ntt.co.jp

Kayo Seki
NTT Advanced Technology Corp.
90-6 Kawakami-cho, Totsuka-ku,
Yokohama-shi, Kanagawa,
244-0805, JAPAN
seki@totsuka.ntt-at.co.jp

Abstract

The effectiveness of using the semantic attributes of verbs has been shown in various kinds of natural processing systems, such as machine translation systems. The addition of an attribute value is, however, time-consuming and must be performed by hand by an expert on the attribute value. In this paper, two methods for efficiently adding verbal semantic attributes to a Japanese-to-English valency transfer dictionary in a machine translation system are proposed and evaluated. One method involves a professional analyst mentally writing down decision-tree-like rules from process images when adding an attribute value to each dictionary entry. The other method involves automatically extracting a decision tree for adding attribute values from dictionary entries with semantic attribute values within a transfer dictionary using the decision tree learning program C5.0. We examine the key factors contributing towards the identification of an attribute value in the entries of the transfer dictionary. The proposed method is also applicable for adding semantic attributes effectively for dictionary entries of a bilingual dictionary within a machine translation system.

1 Introduction

Various machine translation systems are in practical use. In each one of these systems, however, the quality of the finished translation results is not satisfactory. This is mainly due to difficulties in limiting the linguistic phenomena handled by machine translation systems. In particular, the analysis of linguistic expressions which require contextual analysis, such as ellipsis and anaphoric reference, is not perfect. Introducing constraints brought about by context requires an enormous volume of knowledge on word meanings that can be used to determine the semantic relationship between one sentence and another sentence.

To avoid an explosion in the required volume of knowledge, techniques have been proposed to classify word meanings and to determine the relationships between words or between sentences using the typical attribute values of the individual words. In particular, in context processing, verbal semantic attributes have become the key factors in analyzing the flow of sentences.

Various efforts have been made at researching verb classifications (Muraki 1985; Tomiura and Yoshida 1986; EDR 1990; Levin 1993), but such efforts still tend to be limited solely to a classification of the semantics of the verbs per se. Research of this type has not sufficiently taken into account the relationships between word meanings and their usage within sentences for a semantic analysis in machine translation, and is not applicable to machine translation systems directly. Consequently, the full benefits of tracking semantic relationships between sentences and eliminating the polysemy of verbs have not been achieved in machine translation.

To overcome this problem, Nakaiwa et al. proposed a system of 106 verbal semantic attributes (VSA) for Japanese verbs; this system considers both the dynamic characteristics of verbs and the relationships of verbs to cases (Nakaiwa et al. 1994).

These attribute values are used to disambiguate the meanings of Japanese and English pattern pairs in a Japanese-to-English valency transfer dictionary; this dictionary contains 16,000 pairs of Japanese valence patterns and their equivalent English syntactic structures (Figure 1) in a Japanese-to-English machine translation system, ALT-J/E (Ikehara et al. 1997). This dictionary consists of pairs of Japanese case-frame patterns derived from Japanese verbs with semantic constraints for their case elements and English patterns corresponding to the Japanese expressions.[1] Maintaining the expressions in the pairs, which indicate common word meanings between the Japanese and English, which differ so vastly in their syntactic structures, enables the elimination of conceptual ambiguities making possible the granting of detailed and accurate attribute values. Therefore, it is possible to disambiguate the meanings of verbs by selecting verbal patterns in a syntactic semantic analysis.

For example, in the case of the pattern in the semantic valence pattern transfer dictionary in Figure 1, if the verb for the Japanese sentence is "*taberu*", and if the meaning of *ga* case (N1) for *taberu* is SUBJECT and that of *o* case (N2) is FOOD, then the Japanese sentence has an English equivalent "N1 eat N2" and the verbal semantic attribute is SUBJ's BODILY ACTION.[2] If the usage of such a pattern can be expressed by a small number of verbal semantic attributes, then it is possible to easily track the semantic relationships of verbs.

When verbal semantic attributes are given to a pair of individual Japanese and English patterns, it is possible to refer to the meanings of the verbs in Japanese as well as in English.

[semantic valence pattern transfer dictionary]
N1(SUBJECTS)-*ga*   N2(FOOD)-*o*   *taberu*   => N1 eat N2   VSA: SUBJ'S BODILY ACTION
eat
[idiomatic expression transfer dictionary]
N1(SUBJECTS)-*wa*   *se-ga*   *takai*   => N1 is tall.   VSA: SUBJ'S ATTRIBUTE
back   high

Figure 1: Japanese to English Valency Transfer Dictionaries
(semantic constraints in parentheses)

At present, however, attribute values are added to patterns in a machine translation system by hand. It therefore takes a lot of time and labor to make a large-scale transfer dictionary that contains verbal semantic attribute values within each dictionary entry with wide coverage (the knowledge acquisition bottleneck). Furthermore, if analysts want to add attribute values to individual patterns, they must be familiar with the attribute system as well as the machine translation system itself. Therefore, should a user of a machine translation system want to make a dictionary entry in the user transfer dictionary for a specific domain, s/he must first ask a professional analyst of the attribute system for help on adding an attribute value for the dictionary entry. It is, however, impossible for a professional analyst to help every end user who wants to make a dictionary entry for his/her translation target domain by adding a correct attribute value for the entry. Because of this problem, it is necessary to create a method of supporting the addition of attribute values even by non-professional analysts or a method of adding attribute values automatically.

To overcome the problem, two types of methods can be used for the addition of a semantic attribute value for an entry. One typical method involves a professional analyst mentally writing down the process images when adding an attribute value for a dictionary entry. Another method involves automatically extracting rules for adding attribute values using dictionary entries with semantic attributes within

---

[1] The idiomatic expression transfer dictionary contains a core sector of idiomatic expressions such as *"Abura o uru"* literally, "to sell oil", but idiomatically, "to idle away time".

[2] In the case of the Japanese verb *taberu,* there are six patterns in the semantic valence pattern transfer dictionary. Depending on the meanings of the cases co-occurring with *taberu,* the machine translation system selects the best English equivalent from among the six patterns.

a transfer dictionary by analyzing the correlation between the type of attribute value and the characteristics of those dictionary entries having this attribute value. The former method is thought to be an efficient and accurate method for value addition because the method directly uses the know-how of expert analysts. Expert analysts, however, require a lot of time to make near perfect lists. In addition, there is a strong possibility of human error.

In the latter method, it is possible to automatically make complete addition rules using a stochastic analysis program or a decision tree learning program. This method can extract addition rules that are difficult for humans to make or rules that can only be extracted by using stochastics. Even in the latter method, however, if the co-occurrence between an attribute value and a dictionary entry is low, the reliability of any rule made by the low frequency correlation also becomes low. For a high accuracy, this method must carefully selects the attribute values to input into the learning program.
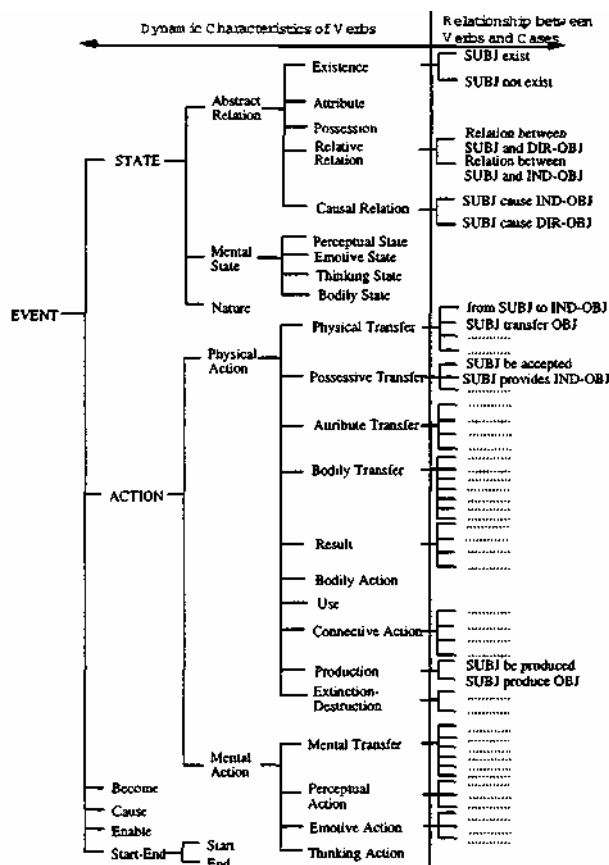

Figure 2: System of Verbal Semantic Attributes

In this paper, we evaluate both methods of adding verbal semantic attributes to a Japanese-to-English valency transfer dictionary in a machine translation system as a first step towards overcoming the known problems in the methods and combining their merits. We examine key factors in identifying attribute values in the entries of the transfer dictionary.

2 A System of Verbal Semantic Attributes

In this section, a system of verbal semantic attributes is described. Nakaiwa et al. (1994) proposed a system of verbal classifications based on the following two factors.

a) **Dynamic Characteristics of Verbs**

This factor is a classification based on a verb's meaning and its effect on the discourse. It is based on the type of action that occurs when the verb is expressed and the situations brought about.

(1) *motsu* "to have" — POSSESSION
kaihatsusuru "to develop" — PRODUCTION

The verb *motsu* "to have" indicates that there is an act of possession within the context.[3] In contrast, the verb *kaihatsusuru* "to develop" indicates that there is

---

[3] There are eleven patterns for *motsu* in our semantic valence pattern transfer dictionary. Among the eleven patterns, the VSAs of four patterns are POSSESSION, those of three patterns are ATTRIBUTE, those of two patterns are EMOTIVE ACTION, that of one pattern is POSSESSIVE TRANSFER, and that of one pattern is NATURE.

something being produced within the context.[4]

b) **Relationship of Verbs to Cases**

This factor is a classification based on the roles which the cases play with the verbs that govern them. It is based on the roles played by the case elements governed by the verbs expressed.

(2) *kanseisuru* "SUBJ be completed" — SUBJ BE PRODUCED
    kaihatsusuru "SUBJ develops OBJ" — SUBJ PRODUCES OBJ

kanseisuru "to complete" and *kaihatsusuru* "to develop" are both verbs indicating acts of production. *kanseisuru* indicates that the SUBJ is being produced, whereas *kaihatsusuru* indicates that the SUBJ produces the OBJ.

The top levels of the system of 106 verbal semantic attributes are shown in Figure 2.

3 Addition of Attribute Values by using a Handmade Decision Tree

The attribute values of the verbal semantic attributes explained in section 2 were originally designed by an analyst who examined each entry of the valency transfer dictionary. Therefore, all of the conditions for determining an attribute addition for a dictionary entry are in the analyst's brain. For rule creation, it seems only natural to write down the meta-rules within the analyst's brain. Moreover, should a person who is not an expert of the attribute system try to add a new attribute for a new dictionary entry, it is effective to allow the person to check the rules based on the information within the dictionary entry.

Two strategies can be used to make rules for attribute value addition by hand:
a) Classify dictionary entries with the same attribute values and extract the common features and/or the typical examples for each attribute value.
b) Extract the decision conditions and the strategies used by an analyst when s/he tries to add a new attribute value for a new dictionary entry, and create the addition flow in a decision tree style.

Rule (a) is suitable for determining which attribute value is the most suitable one from among a few attribute value candidates for a specific dictionary entry. It is not suitable, however, for someone without knowledge on the system to determine an attribute value from among all possible attribute values. Rule (b), in contrast, is suitable for a person lacking expertise to determine an attribute value because the attribute value will finally be decided on only by the selection of one answer at each query step within decision-tree-like rules. Rule (b) has some problems, however, because the number of queries is limited due to the limited number of human judgments and also because there is no guarantee on the suitability of the queries within the decision tree as they involve handmade rules.

To overcome these problems, we use the following two rules for this examination; a) a table explaining the definition and examples of each attribute value, and b) decision-tree-like rules for determining an attribute value designed by one expert analyst (Figure 3).

The method of adding an attribute value for a new dictionary entry using the decision-tree-like rules is done with the following steps.

Step 1 Judge the part of speech of the Japanese verb in an entry.
Step 2 Examine the meaning of the Japanese verb in the entry by referring to the whole transfer pattern information.
Step 3 Select an answer for a query within the decision tree.
Step 4 Examine an equivalent expression for the Japanese verb or refer to a different pattern with the same Japanese verb if the answer for a query in the decision tree is not suitable; select the most suitable answer.

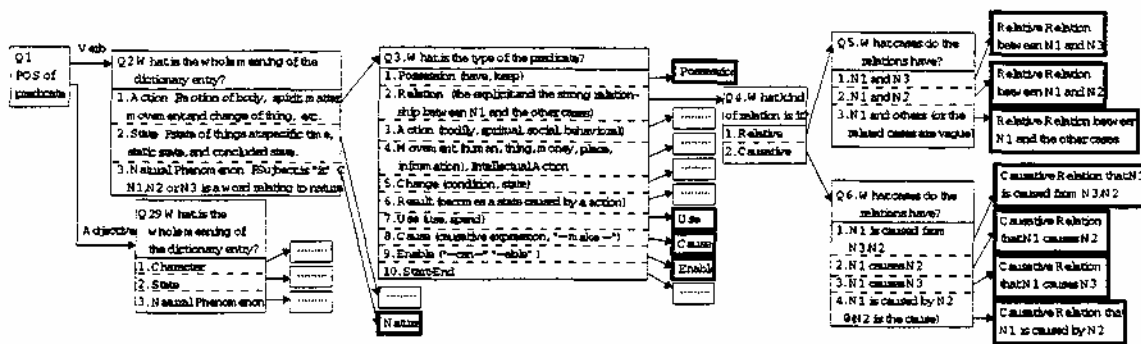[4] There is only one pattern for *kaihatsusuru* whose VSA is PRODUCTION in our dictionary.

Figure 3: Handmade Decision Tree

4 Addition of Attribute Values by using a Program-made Decision Tree

In contrast to section 3, we examine a method of automatically creating a decision tree for attribute value addition for a dictionary entry by using the stochastic characteristics between an attribute value and the feature of a dictionary entry. A decision tree learning program C5.0 (Quinlan 1998), which was originally based on the decision tree algorithm ID3 (Quinlan 1986), is used to automatically make the decision tree.

C5.0 makes a decision tree from the list of feature vectors extracted from each dictionary entry for the selection conditions of a query within the decision tree and an attribute value.

4.1 Feature Vectors

The performance of a decision tree made by a decision tree learning program highly depends on the kind of feature vector used as the input to the program. In the following subsections, we explain the feature vectors used to make rules for adding a verbal semantic attribute value for a dictionary entry in the Japanese-to-English valency transfer dictionary shown in Figure 1.

4.1.1 Type of Valency Pattern (Case-roles)

Verbal semantic attributes are designed based on the meaning of a verb and the relationship between the case elements and the verb. Accordingly, we determine an attribute value using the types of case relationships. In the entries to the ALT-J/E transfer dictionary (Ikehara et al. 1997), the case relationships are each labeled as a variable of cases: N1 (subject; mainly the *ga*-case in Japanese), N2 (direct object; mainly the *o*-case in Japanese), N3 (indirect object; mainly the *ni*-case in Japanese), and so on. Some verbal semantic attributes are defined according to the case relationships (for example, RELATIVE RELATIONSHIP BETWEEN N1 AND N3). For this reason, the labels of the case relationships within a dictionary entry are used as a feature vector.

4.1.2 Semantic Constraints for Cases within a Japanese Pattern

Some attribute values are defined according to the types of semantic constraints for cases within a Japanese pattern; for example, N1's (HUMAN/ANIMAL) BODILY ACTION and N1 (SUBJECT) USES N2. Therefore, such semantic constraints are also used as feature vectors.

4.1.3 English Words within an English Pattern

Each dictionary entry in the Japanese-to-English valency transfer dictionary consists of a Japanese pattern part and an equivalent English pattern part. Therefore, even if the Japanese verb within a Japanese pattern part is ambiguous, it is possible to disambiguate the meaning of the Japanese verb by pairing the Japanese verb with its equivalent English verb (Nakaiwa et al. 1994). As a result of this characteristic of the dictionary, the types of English words within an English pattern part are effective for determining the attribute value for a dictionary entry. This characteristic is also

189

provided when the relationship between an attribute value and the information of each dictionary entry is analyzed to design a handmade decision tree (Figure 3). For example, if an English pattern contains an English word "walk" or "eat", then the attribute value of the dictionary entry tends to be N1's (HUMAN/ANIMAL) BODILY ACTION.

4.1.4   Semantic Category of a Japanese Verb

An attribute value classified based on the dynamic characteristics of a verb (left part of Figure 2) has a strong correlation with the semantic category of the Japanese verb. For example, the verbal semantic attribute BODILY ACTION contains many patterns whose Japanese verbs have a specific semantic category, OPERATION, such as *nageru* "to throw" and *untensuru* "to drive". The semantic category of a Japanese verb within a Japanese pattern is used here.

4.2   Process for Learning a Decision Tree

The process for learning a decision tree from the Japanese-to-English valency transfer dictionary whose dictionary entries have verbal semantic attributes is summarized in the following.

Step 1 Extraction of pairs of a feature vector and an attribute value.
From each entry in the Japanese-to-English valency transfer dictionary like in Figure 1, pairs of a feature vector consisting of the features explained in section 4.1 and an attribute value are extracted.

Step 2 Modification of semantic categories in a feature vector.
Those semantic categories extracted as a feature vector in step 1, are from some 2,718 categories. Each semantic category is connected with tree-structured links (Ikehara et al. 1997). For example, the semantic category OPERATION has hypernym categories: BUSINESS LABOR, LABOR, ACTION, HUMAN ACTIVITY, THING, ABSTRACT, and NOUN. Therefore, even if a feature vector contains the category OPERATION, there is no guarantee that OPERATION will be suitable as a condition for determining an attribute value or whether one of its hypernym categories is more suitable as the condition. To overcome this problem, we select feature vectors from not only the semantic category in a dictionary entry itself but also from its hypernym categories, by using the encoding method proposed by Almuallim et al. (1994).

Step 3 Modification of feature vectors to be a format for C5.0 input.
Feature vectors extracted in the previous steps are modified to a format for C5.0 input. Actually, the number of occurrences of each feature within all of the feature vectors extracted from the valency transfer dictionary is counted. Only those features whose occurrence count is equal to or more than a threshold value are used as feature vectors in C5.0 after the above modification to the C5.0 input binary format, indicating whether or not each feature is contained in a dictionary entry.

Step 4 The input of feature vectors into C5.0 and the creation of a decision tree.
Those feature vectors modified in step 3 are input into the C5.0 program as training data. Then, a decision tree that reflects the characteristics within the training data is generated.

Figure 4.2 shows an example of a program-made decision tree. In the tree, "2079 = 1" indicates a query, i.e., whether the category of a verb in a dictionary entry is "2079" (DOWNFALL) or not. If the category is "2079" and if the dictionary entry does not contain case "N2", then the attribute value of the entry is judged as "(32 1)" (SUBJ CHANGES STATE RELATIVE TO TIME) (N2 = 0: (32 1) (7.0)); if the dictionary entry contains case "N2", the attribute value is judged as "(32 2)" (SUBJ CHANGES STATE WITHOUT INTENTION) (N2 = 1: (32 2) (3.0)). If the category of a verb in a dictionary entry is not "2079", the following queries are examined and finally an attribute value is automatically identified depending on the characteristics of the dictionary entry.

5 Evaluation

The handmade decision tree explained in section 3 and the program-made decision tree explained in section 4 are compared for the accuracy of the addition of a verbal semantic attribute for an entry to the valency transfer dictionary in ALT-J/E (Ikehara et al. 1997).

5.1 Evaluation
of the Handmade Decision Tree

The handmade decision tree shown in Figure 3 was evaluated by examining whether or not a non-expert can easily and accurately add a verbal semantic attribute using the handmade decision tree.

5.1.1 Evaluation Method

The conditions for the evaluation were as follows.

```
2079 = 1:
:...N2 = 0: (32 1) (7.0)
:   N2 = 1: (32 2) (3.0)
2079 = 0:
:...1871 = 1: (5 0) (12.0/2.0)
    1871 = 0:
    :...."live" = 1:
        :...1593 = 0: (1 0) (15.0/5.0)
        :   1593 = 1: (3 1) (3.0)
        "live" = 0:
        :..."win" = 1:
            :...2065 = 1: (16 1) (2.0/1.0)
            :   2065 = 0:
            :   :...1705 = 0: (31 6) (15.0)
            :       1705 = 1: (31 3) (3.0/1.0)
            "win" = 0:
            :..."hear" = 1:
                :...2178 = 1: (22 7) (2.0)
                :   2178 = 0:
                :   :...N2 = 0: (8 3) (2.0/1.0)
                :       N2 = 1: (23 2) (6.0)
                "hear" = 0:
                :...2364 = 1: (11 0) (10.0)
                    2364 = 0:
                    :..............
```

Figure 4: Program-made Decision Tree

Target of Attribute Addition Verbal semantic attributes were added for 100 randomly selected entries with verbal semantic attributes of the Japanese-to-English valency transfer dictionary shown in Figure 1 in a Japanese-to-English machine translation system, ALT-J/E (Ikehara et al. 1997). Seventeen entries out of the 100 entries had more than one attribute value.

Attribute Adder (Tester) The attribute values were added by the following two non-experts on verbal semantic attributes; (A) a person familiar with the Japanese-to-English valency transfer dictionary but not familiar with verbal semantic attributes, (B) a person familiar with neither the Japanese-to-English valency transfer dictionary nor verbal semantic attributes. The results of the two persons were compared to evaluate the difference in addition accuracy depending on the familiarity of the Japanese-to-English valency transfer dictionary.

Reference Materials for Attribute Adder Four materials were used when the non-experts added attribute values:

- The hand made decision tree in Figure 3
- A table explaining the definition and examples of each attribute value
- A list of dictionary entries for adding attribute values (100 pattern pairs)
- Dictionaries (Japanese dictionary and Japanese-English dictionary)

Attribute Addition Process After receiving a short talk (overview) about the verbal semantic attribute system and the Japanese-to-English valency transfer dictionary, the non-experts added attribute values for five dictionary entries as training. After that, the non-experts added attributes for the 100 dictionary entries. We did not announce any limitations on the number of attribute values for one dictionary to the non-experts.

Successfully Added Dictionary Entries When the non-experts added an identical verbal semantic attribute value which had already been added for a dictionary entry, or when they added a different attribute value but the added attribute value was also acceptable for the dictionary entry according to an examination by the expert, the dictionary entry was judged to be successfully added.

191

5.1.2 Evaluation Results

The addition accuracy using the handmade decision tree by the non-experts is shown in Table 1. The accuracy of A's added attribute value was higher than that of B. In addition, the addition time for A's 100 dictionary entries was shorter than that of B. These results indicate that the handmade decision tree in Figure 3 did not allow to non-experts on both the verbal semantic attributes and the Japanese-to-English valency transfer dictionary to achieve a high accuracy. Non-expert A, however, who was not an expert on verbal semantic attributes, could correctly add 78% for the dictionary entries. This result indicates that the decision tree can provide enough information for an analyst who makes new dictionary entries to add verbal semantic attributes.

In the results of a detailed examination, it was found that the two non-experts added the same attribute values at a 22% rate for the dictionary entries. Those attribute values showed some specific attribute values such as ATTRIBUTE, PHYSICAL TRANSFER, and BODILY ACTION. This result shows that the rules for adding attribute values within the handmade decision tree are well-designed; in other words, these attribute values are easy to understand even for non-experts. Furthermore, we classified those dictionary entries whose attribute values could not be added correctly by the non-experts into four types according to an examination of these entries.

(a) **Misunderstanding of the meaning** - A: 9 items, B: 32 items
There was a big difference between the results of A and B. This was mainly due to the different knowledge about the dictionary. To minimize the difference, support tools are needed to help non-experts understand the meaning of each dictionary entry.

(b) **Unsuitability of the decision tree** - A: 4 items, B: 11 items
Some rules within the decision tree did not have enough information or were difficult to understand for an attribute value decision. Errors can be minimized by analyzing the types of errors that occur and by adding examples and explanations.

(c) **Lack of information within a dictionary entry** - A: 3 items, B: 3 items
This were errors caused by individual dictionary entries with no semantic constraints. Dictionary entries needed to be modified to overcome these errors.

(d) **Addition errors at the level of the relationships of verbs to cases** - A: 5 times, B: 6 times
These errors were caused when a category of the dynamic characteristics of a verb (left part of Figure 2) was correct but a category of relationships of verbs to cases (right part of Figure 2) was incorrect. Documents and support tools need to be modified to clarify the difference of each category in the relationships of verbs to cases.

The results also showed that 23-25% of correctly added attribute values were not the same as already added attribute values but are acceptable (18 items out of 78 items for A and 11 items out of 44 items for B). This result indicates that these patterns each had more than one attribute candidate depending on the point of view even by the addition of an expert on the attribute system. Consequently, about 80% is the theoretical upper limit for adding the same attribute values as already added attribute values.

Table 1: Addition Accuracy of Attribute Values using a Handmade Decision Tree

Attribute adder	Addition accuracy		Addition time
A (dictionary expert)	78 %		55 minutes
	same: 60 %	acceptable: 18 %	
B (dictionary non-expert)	44 %		83 minutes
	same: 33 %	acceptable: 11 %	

5.2 Evaluation of the Program-made Decision Tree

We evaluated the program-made decision tree automatically created by the decision tree learning program C5.0 (Quinlan 1998).

5.2.1 Evaluation Method

The conditions for the evaluation were as follows.

Selected Feature Vectors and Dictionary Entries for Decision Tree Learning
Dictionary entries whose verbal semantic attributes have already been added to the Japanese-to-English valency transfer dictionary were used as the training data for C5.0. The following four kinds of information within a dictionary entry were used as feature vectors for C5.0. (section 4.1)

(a) Type of Valency Pattern (used in all conditions)
(b) Types of English Words within an English Pattern
(c) Semantic Constraints for Cases within a Japanese Pattern (patterns with only N1 or with only N1 and N2)
(d) Semantic Category of Japanese Verb

The addition accuracy was examined in terms the employed feature vectors (Table 2).

We did not include dictionary entries that added more than one attribute value in the training data. Dictionary entries that contained selected feature vectors were used as the training data. For example, in an examination of semantic constraints for cases, just dictionary entries with only N1 (3,748 entries) or dictionary entries with only N1 and N2 (3,130 entries) were used as the training data.

Execution Parameter of the Decision Tree Learning Program We carried out C5.0 decision tree learning without setting any special parameters. The accuracy was also examined in terms of the threshold value (explained in step 3 of section 4.2).

Target of Attribute Addition Those dictionary entries used in the decision tree learning were also used as the target dictionary entries for the attribute addition. The same evaluation conditions as in the evaluation of the handmade decision tree were kept for the entries because the handmade decision tree was created by having an expert examine all of the dictionary entries.

Successfully Added Dictionary Entries When the automatically created decision tree could add the same verbal semantic attribute already added for each dictionary entry, the dictionary entry was judged to be successfully added.

5.2.2 Evaluation Results

The addition accuracy of the program-made decision tree, which depended on the selected feature vectors used, is shown in Table 3. In the evaluation, the threshold value was 1. As shown in the table, when (b) English words and (d) the category of the Japanese verb were used, the accuracy of the added attribute value achieved the highest value (70.4%). The accuracy did not increase more than the accuracy obtained in only using (b) and (d), even when (b), (d), and (c) constraints for the cases were used as feature vectors. This was due to the data sparseness for the features of the

Table 2: Selected Feature Vectors and Dictionary Entries for Decision Tree Learning

(a) Valence pattern	Selected feature vectors			The number of dictionary entries
	(b) English words	(c) Constraints for cases	(d) Category of Japanese verb	
used	unused	unused	used	12,601
	used	unused	unused	12,601
	used	unused	used	12,601
	used	used (N1)	used	3,748
	used	used (N1, N2)	used	3,130

193

semantic constraints for the cases and the lack of adequate information on the semantic constraints for the cases.

Table 4 shows the addition accuracy in terms of the threshold value for (b) and (d) when setting the selected feature vectors to obtain the best feature result for (b) and (d). As shown in the table, even when the threshold value for (b) and (d) was increased to two and the dimension of the input feature vectors became about 60% of the dimension of inputted vectors when the threshold value for (b) and (d) was one, the accuracy was the same (70.4%). Furthermore, when the threshold value for (b) was increased, both the dimension of the input feature vectors and the number of rules in the decision tree drastically decreased. When the ratio (b):(d) was decreased from 2:5 to 10:5, the cpu time also decreased to about 12%. However, even when the threshold value was increased, the addition accuracy did not decrease very much as when (b) and/or (d) were unused; (b) unused: 59.8%; (d) unused: 61.0%; threshold of (b):(d)=10:5: 65.4%. These results indicate that (b) and (d) were better, and depending on the required cpu power, the threshold value for (b) and (d) could be selected to make an effective and efficient decision tree.

Table 3: Addition Accuracy for Selected Feature Vectors

Selected feature vectors			Dimension of the input feature vectors	The No. of rules in the decision tree	Addition accuracy
(b) English words	(c) Constraints for cases	(d) Category of Japanese verb			
unused	unused	used	1,531	1,252	59.8%
used	unused	unused	4,677	1,706	61.0%
used	unused	used	6,190	2,004	70.4%
used	used (N1)	used	3,341	1,171	68.9%
used	used (N1, N2)	used	2,870	993	68.3%

Table 4: Addition Accuracy for Threshold Value

Threshold value		Dimension of the input feature vectors	The No. of rules in the decision tree	Addition accuracy
(b) English words	(d) Category of Japanese verb			
1	1	6,190	2,004	70.4%
2	1	4,021	2,004	70.4%
2	2	3,847	2,004	70.4%
2	5	3,603	1,990	70.3%
5	1	2,240	1,783	67.7%
5	5	2,022	1,794	67.8%
10	5	1,460	1,640	65.4%

5.3 Combination of Two Methods

In this section, we examine the addition accuracy of the handmade decision tree (section 5.1) and the program-made decision tree (section 5.2) and propose how to combine these two method and achieve a high accuracy. In the case of attribute addition using the handmade decision tree, in spite of the attribute addition work by humans, the addition accuracy by a non-expert not familiar with the dictionary (B) only reached 44%, but the addition accuracy by a non-expert familiar with the dictionary (A) reached as high as 78%. In the case of the program-made decision tree, in contract, even with automatic addition, the accuracy was reached 70.4%. This comparison is not fair for the result of the program-made decision tree because successfully added entries for the program-made decision tree did not contain entries whose added attribute values were acceptable. In fact, when the same condition were applied for successfully added entries for the result of the handmade decision tree, the accuracy of B was only 33% and even A was 60%; this result shows that the method using the program-made decision tree already achieves a higher accuracy than the method of addition by a non-expert using the handmade decision tree.

To achieve a higher accuracy, we examined the results of automatic addition by the program-made decision tree for those dictionary entries used for the evaluation of the handmade decision tree. The decision tree which achieved the best result (70.4%;

threshold of (b):(d)=l:l in Table 4) was applied for 77 entries which had only one already added attribute value and whose Japanese verb had a category in the 100 entries used for the evaluation of the handmade decision tree. According to the results, the same attribute value with an already added attribute value was added for 51 items out of the 77 items (66%) and an acceptable attribute value was added for 8 items (10%). This result indicates that almost the same accuracy, 77%, as the result by A was achieved for the 77 entries. According to a further examination of the failed 18 items by the program-made decision tree, the added attribute values of 9 items were rejected if the obligatory condition depending on the attribute value within the handmade decision tree was directly applied. This result indicates that to achieve a combined method, the use of the obligatory condition within the handmade decision tree is effective for achieving a more accurate method.

6   Conclusion

This paper has examined two methods for the addition of verbal semantic attributes to a Japanese-to-English valency transfer dictionary in a machine translation system. One method uses a handmade decision tree designed by an expert on verbal semantic attributes. The other method uses a program-made decision tree designed by a decision tree learning program extracted from dictionary entries with attribute values. From a comparison of the two methods, we found that the combination of the two methods appears to be more effective in achieving a higher accuracy.

In the future, we plan to perform a detailed evaluation of these two methods and effectively combine the two methods. Furthermore, we plan to examine a user-friendly and efficient attribute addition tool for human interaction purposes. We will also examine the combination of the proposed method with other methods involving word sense disambiguation using statistics within a corpus (Yarowsky 1992; Almuallim et al. 1994; Tanaka 1994; Ide et al. 1998).

References

Almuallim, Hussain. et al.: 1994, 'Two Methods for Learning ALT-J/E Translation Rules from Examples and a Semantic Hierarchy', in *15th International Conference on Computational Linguistics: COLING-94*, Kyoto, pp. 57-63.

EDR: 1990, *Concept Dictionary,* TR-027, EDR.

Ide, Nancy & Jean Véronis: 1998, 'Introduction of the Special Issue on Word Sense Disambiguation: The Stare of the Art', in *Computational Linguistics,* Vol. 24, No. 1, pp. 1-40.

Ikehara, Satoru, et al. eds.: 1997, *Nihon-go Goi-Taikei [Goi-Taikei - A Japanese Lexicon],* Iwanami Publishing Co (in Japanese).

Levin, Beth: 1993, *English Verb Classes and Alternations,* The University of Chicago Press.

Muraki, Shinjiro: 1985, 'Jyutsugo-so niyoru dousi no bunrui [Classification of Verbs by Predicates]', in *SIG notes of Information Processing Society of Japan,* Vol. NL-48 No. 5 (in Japanese).

Nakaiwa, Hiromi, Akio Yokoo & Satoru Ikehara: 1994, 'A System of Verbal Semantic Attributes Focused on the Syntactic Correspondence between Japanese and English', in *15th International Conference on Computational Linguistics: COLING-94,* Kyoto, pp. 672-678.

Quinlan, J. Ross.: 1986, 'Induction of Decision Trees', in *Machine Learning,* Vol. 1, No. 1, pp. 81-106.

Quinlan, J. Ross.: 1998, http://www.rulequest.com/.

Tanaka, Hideki: 1994, 'Verbal Case Frame Acquisition from a Bilingual Corpus: Gradual Knowledge Acquisition', in *15th International Conference on Computational Linguistics: COLING-94,* Kyoto, pp. 727-731.

Tomiura, Youichi. and Sho Yoshida: 1986 'A Research of the Polysemy and Description of Verbs', in *SIG notes of Information Processing Society of Japan,* Vol. NL-55 No. 2 (in Japanese).

Yarowsky, David: 1992, 'Word-Sense Disambiguation Using Statistical Methods of Roget's Categories Trained on Large Corpora', *14th International Conference on Computational Linguistics: COLING-92,* Nantes, pp. 454-460.