The Effects of Negative Premises on Inductive Reasoning: A Psychological Experiment and Computational Modeling Study

Kayo Sakamoto (SAKAMOTO@Nm.Hum.Titech.Ac.Jp)
Tokyo Institute of Technology, 2-21-1 O-okayama,
Meguro-ku, Tokyo, 152-8552 JAPAN

Masanori Nakagawa (NAKAGAWA@Nm.Hum.Titech.Ac.Jp)
Tokyo Institute of Technology, 2-21-1 O-okayama,
Meguro-ku, Tokyo, 152-8552 JAPAN

Abstract

Various learning theories stress the importance of negative learning (e.g., Bruner, 1959; Hanson, 1956). However, the effects of negative premises have rarely been discussed in any detail within theories of inductive reasoning (with the exception of Osherson et al., 1990). Although Sakamoto et al. (2005) have proposed some computational models that can cope with negative premises and verified their psychological validity, they did not consider cases where category-based induction theory is ineffective, such as when the entities in both negative and positive premises belong to the same category. The present study was conducted to test the hypothesis that, even when negative and positive premises involve same-category entities, people can estimate the likeliness of an argument conclusion by comparing feature similarities. Based on this hypothesis, two computational models are proposed to simulate this cognitive mechanism. While both these models were able to simulate the results obtained from the psychological experiment, a perceptron model could not. Finally, we argue that the mathematical equivalence (from Support Vector Machines perspective) of these two models suggests that they represent a promising approach to modeling the effects of negative premises, and, thus, to fully handling the complexities of feature-based induction on neural networks.

Introduction

This study is concerned with evaluating "arguments", such as:

Collies produce phagocytes.
Horses produce phagocytes.

Shepherds produce phagocytes.

The propositions above the line are referred to as "premises" while the statement below is the "conclusion". The evaluation of an argument involves estimating the likelihood of the conclusion based on the premises. Osherson, Smith, Wilkie, Lopez, and Shafir (1990) refer to this kind of argument as a "categorical" argument, because the predicate (e. g., "produce phagocytes") in the premises and conclusion is attributed to one or more entities (e. g., "Collies", "Shepherds").

The premises can also be negative in form (e. g., "Penguins do not produce phagocytes"). Since classic studies, such as discrimination learning (e. g., Hanson, 1956) and concept learning (e.g., Bruner,1959), the importance of negative examples has been widely recognized, and has been demonstrated in more recent studies, such as causal learning (e.g., Buehner and Cheng, 2005). However, the effects of negative premises have rarely been discussed in any detail in the context of inductive reasoning studies concerned with the evaluation of arguments.

Investigating the effects of negative premises can undoubtedly contribute to our understanding of inductive reasoning. For instance, cases where the entities in both negative and positive premises belong to the same category are clearly problematic for the category-based induction theory (Osherson et al., 1990) because it is impossible to distinguish between negative premises and positive premises from the categorical viewpoint. In contrast to the similarity and coverage model based on category-based induction theory, Sloman (1993) has proposed a feature-based model based on a simple perceptron. According to Sloman, knowledge of category structure is not required for the evaluation of arguments. Rather, he assumes that argument evaluation is based on a simple computation of feature similarities between the entities of the premises and the conclusion. Thus, Sloman's feature-based model may be more effective at coping with negative premises than category-based induction theory. However, the psychological validity of Sloman's model has yet to be tested in terms of processing negative premises. Accordingly, this study examines the validity of feature-based induction theory to handle these cases that are so problematic for category-based induction theory.

In terms of model construction, what kind of model can adequately represent the cognitive process of feature-based induction, including negative premises? In addition to Sloman's (1993) model, Sakamoto, Terai and Nakagawa (2005) have also proposed a feature-based model. While structurally similar, their model extends the learning algorithm in order to cope with negative premises. Moreover, their model utilizes corpus-analysis results to compute feature similarities, rather than the results of psychological evaluations used in Sloman's model. This means that Sakamoto et al's model is capable of simulating a far greater variety of entities than Sloman's model (over 20,000 compared to just 46), because the corpus analysis provides information for an enormous quantity of words. However, when induction of the appropriate category is difficult, then the level of computation involved in the feature similarity comparisons will far exceed the computational capacity of simple perceptrons. This study therefore proposes a modified version of the Sakamoto et al model, referred to as the "multi-layer feature-based neural network model", which is compared with the previous perceptron model.

A further alternative model for coping with negative premises would also seem to be possible. Osherson, Stern, Wilkie, Stob, and Smith (1991) have also proposed a model to handle negative premises based on feature similarity, involving more complex computation than possible with perceptrons. However, the Osherson et al.'s feature-based model requires knowledge of relevant taxonomical categories, and also utilizes psychological evaluations like Sloman's model. As the restricted number of available entities (46, similar to Sloman's model) makes it difficult to apply that model, this study also modifies the Osherson et al feature-based model in order to handle greater numbers of entities and to eliminate the need for categorical knowledge.

The outline of this study is as follows: First, the corpus analysis is described. The results of the corpus analysis were utilized in creating clear category definitions and in constructing the models. Second, a psychological experiment is described which was conducted to examine the effects of negative premises, that cannot be accounted for by the category-based induction theory. Third, two models—the multilayer feature-based model and a modified version of Osherson et al's feature-based model—are proposed and tested in terms of their psychological validity. Finally, this study argues that the mathematical equivalence (from a support vector machine perspective) of these two models suggests that they represent a promising approach to modeling the complexities of feature-based induction on neural networks.

A Corpus Analysis for Category Definitions and Model Construction

Categories used in this study are defined as latent semantic classes estimated from an analysis of the words in a Japanese corpus. The estimations were based on a soft-clustering of words according to modifying frequencies in the corpus. The soft-clustering results are represented as conditional probabilities of words given the latent classes. From these probabilities, the conditional probabilities of feature word/phrases, given particular nouns, are also computed. The models in this study applied these conditional probabilities of feature words as the strengths of the relationships between nouns (entities) and features.

The method of soft-clustering was based on a method of similar structure to Pereira's method or PLSI (Hofmann 1999; Kameya & Sato 2005; Pereira, Tishby, and Lee 1993). This method assumes that the co-occurrence probability of a term "N_i" and a term "A_j", $P(N_i,A_j)$, can be represented as formula (1):

$$P(N_i, A_j) = \sum_k P(N_i \mid C_k) P(A_j \mid C_k) P(C_k), \quad (1)$$

where $P(N_i|C_k)$ is the conditional probability of term N_i, given the latent semantic class C_k. Each of the probabilistic parameters in the model, $P(C_k)$, $P(N_i|C_k)$, and $P(A_j|C_k)$ are estimated as values that maximize the likelihood of co-occurrence data measured from a corpus using the EM algorithm (See Kameya & Sato, 2005). In this study, term "N_i" represented a noun, and term "A_j" represents a feature word, such as a predicate. The number of latent classes was fixed at 200.

For the actual estimation, the word co-occurrence frequencies used were extracted from Japanese newspaper articles, covering a ten-year span (1993-2002) of the Mainichi Shimbun. This co-occurrence frequency data comprises the combinations of 21,205 nouns and 83,176 predicates in modification relations. CaboCha (Kudo & Matsumoto, 2002), a Japanese analysis tool for modification relations, was used for extraction.

In order to test the assumption that the latent classes correspond to categories, it is important to identify the meanings of the classes. In this case, it is possible to identify the meaning of a class from the conditional probability of the latent class C_k given a predicate A_j ($P(C_k|A_j)$) and the probability of the latent class C_k given a noun N_i ($P(C_k|N_i)$), which can be computed from the estimated parameters $P(C_k)$, $P(N_i|C_k)$, and $P(A_j|C_k)$, applying the Bayesian theorem.

Class of Foods(c1)

	P(c1|ni)		P(c1|aj)	
1	steak	0.876	boil down	0.987
2	set meal	0.867	eat	0.978
3	grain foods	0.811	fill one's mouth with	0.937
4	vegetable soup	0.817	want to eat	0.927
5	meat	0.739	stew	0.920
6	curry	0.734	don't eat	0.918
7	Chinese noodle	0.720	can eat	0.913
8	pizza	0.716	boil	0.894
9	barleycorn	0.594	clear one's plate	0.894
10	rice cake	0.555	let's eat	0.892

Class of Valuable assets(c2)

	P(c2|ni)		P(c2|aj)	
1	stock	0.929	issue	0.929
2	government bonds	0.862	list	0.916
3	place	0.791	increase	0.899
4	building estate	0.780	release	0.884
5	real estate	0.757	vend out	0.877
6	cruiser	0.662	sell	0.852
7	farmland	0.657	borrow on	0.841
8	foreign bonds	0.628	not sell	0.802
9	house	0.594	buy and add	0.802
10	currency	0.555	keep	0.781

Table 1.
Examples of estimated classes and their representative members

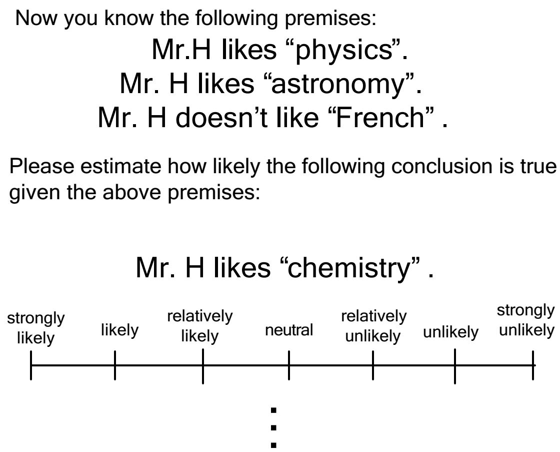

Figure 1. Example of inductive reasoning tasks

This probability denotes the class membership of each word. Based on the estimations, most of the latent classes were identified as meaningful categories, as shown in Table 1.

From the estimated parameters $P(C_k)$, $P(N_i|C_k)$, and $P(A_j|C_k)$, it is also possible to compute the conditional probabilities of feature words given particular nouns, as follows:

$$P(A_j \mid N_i) = \frac{\sum_k P(A_j \mid C_k) P(N_i \mid C_k) P(C_k)}{\sum_k P(N_i \mid C_k) P(C_k)}. \quad (2)$$

In this study, this conditional probability $P(A_j \mid N_i)$ is assumed as the strengths of the relationships between features and entities. When a certain feature word has a high conditional probability given a particular noun, it is natural that the entity denoted by the noun has the feature indicated by the feature word. This conditional probability was therefore applied in the models.

Problematic Experimental Data for the Category-based Induction Theory

This study hypothesizes that people are able to estimate the likeliness of an argument conclusion, even when the entities of both negative and positive premises belong to the same category, by comparing feature similarities. This is something which the usual category-based induction theory cannot fully explain. In order to test this hypothesis, the following psychological experiment was conducted.

METHOD

Participants: Undergraduate students (N = 114) were randomly assigned to one of two inductive reasoning tasks presented in a questionnaire format; 59 students completed one task list, while 55 students completed the other task list.

Materials and Procedure: The questionnaire task lists required inductive reasoning. Each list consisted of three task sets of inductive reasoning arguments. Each set contained two positive premises where the entities belong to a particular category, a negative premise, and thirty conclusions that share the same premise statements (See Figure 1). In the within-category condition, the negative premise belongs to the same category as the positive premises. In contrast, in the between-category condition, the negative premise belongs to a different category from the positive premises (See Table 2).

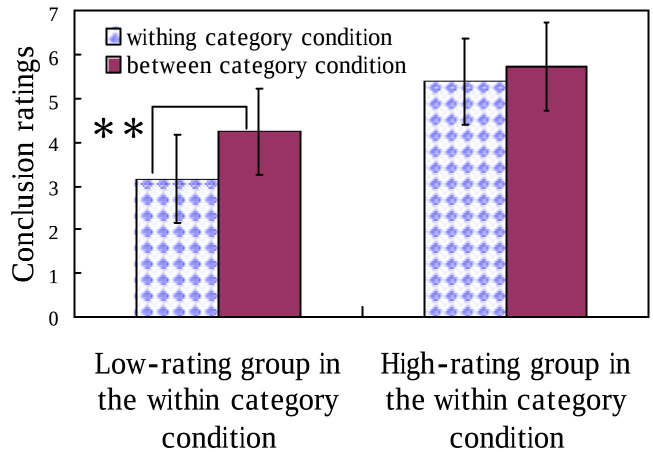

Figure 2. Conclusion likelihood ratings as a function of the within-category and between-category condition, and two rating groups (**: p<0.01)

Table 2. Example of the inductive reasoning task sets (inductive reasoning about the category of learning subject)

	within	**between**
positive premise	"physics"	"physics"
	"astronomy"	"astronomy"
negative premise	"French"	"shopping" (from the category of leisure)
conclusion	chosen from the category of learning subjects	chosen from the category of learning subjects and leisure

Table 3. Example of two rating groups based on the ratings in the within-category condition.

		ratings within	**ratings between**	***P('learning subject'|n)***
High-rating group	"mathematics"	6.000	5.864	0.727
	"arithmetic"	5.745	5.661	0.711
	"chemistry"	5.236	5.492	0.677
	"pharmacy"	4.655	4.712	0.701
Low-rating group	"Japanese literature"	3.836	3.729	0.717
	"English"	3.036	4.136	0.699
	"Chinese"	2.964	3.441	0.801
	"Hangeul"	2.800	3.407	0.701

The premise and conclusion statements all consisted of a combination of a predicate (*Mr. H likes '~'*) and an entity (*curry*), such as *"Mr. H likes curry."* In the case of negative premises, the predicate involved a negative verbal form, such as *"Mr. H doesn't like sports."* As shown in Figure 1, the participants were asked to rate the likelihood of the conclusions on a 7-point scale, given a set of three premises presented above the conclusions. For each task list, the two conditions were counterbalanced among the three task sets.

RESULTS

The results for the two conditions were compared in terms of conclusion ratings. These conclusions were divided equally into a high-rating group and a low-rating group based on the rating scores in the within-category condition. Table 3 shows the members of each group. As shown in Figure 2, only the low-rating group of the within category condition was significantly different from that group of the between category condition. On the other hand, the high-rating group of the within category condition is not strongly different from the group of the between category condition. For example, when the negative premise entity is "French" that belongs to the category of learning subject, the rating of conclusion entity "English" is significantly lower than when the negative premise is "shopping". However, the ratings of conclusion entity "chemistry" in the cases of negative premise "French" and negative premise "shopping," are not radically different. This would suggest that "English" was judged as being similar to "French", and hence its rating with the negative premise "French" was lower than that with the negative premise "shopping". As shown in Table 3, membership to the relevant category ($P(C|N)$) does not in itself yield a simple explanation of the similarity between negative premises and conclusions, which is problematic for the category-based induction theory. In the next section, we will explore a solution to such complex similarity judgments within inductive reasoning by constructing some models based on the feature-based induction theory.

Construction of the Models

Previous Perceptron Model: This study proposes two types of models. However, because their psychological validity is compared with the validity of the feature-based perceptron model proposed by Sakamoto et al. (2005), it is appropriate to start with a brief description of the feature-based perceptron model.

The feature-based perceptron model is an extended version of Sloman's model (Sloman, 1993), which consists of an input layer and one output node, where the weights between the input nodes and the output node are estimated by the usual delta method using the features strengths of positive and negative premises, which are computed as the conditional probability $P(A_j|N_i)$, as previously detailed, according to the following formulas:

$$W(N_i^+) = W(N_{i-1}^+) + [1 - O(N_i^+)]O(N_i^+), \quad (3)$$

$$W(N_i^-) = W(N_{i-1}^-) + [0 - O(N_i^-)]O(N_i^-), \quad (4)$$

$$W(N_i^c) = W(N_{i-1}^-) + [T_i - O(N_i^-)]O(N_i^-), \quad (5)$$

$$O(N_i) = \frac{W \cdot I(N_i)}{|I(N_i)|^2},$$

$$N_i \in N_i^+, N_i^-, N_i^c, \quad (6)$$

where N_i^+ is the ith positive premise entity, N_i^- is the ith negative premise entity, and N_i^c is the ith conclusion entity. $W(N_i^+)$, $W(N_i^-)$ and $W(N_i^c)$ indicate the weights when N_i^+, N_i^- and N_i^c are encoded as the premises and the conclusion, respectively. T_i denotes the target value for the ith conclusion. This value is obtained from psychological experiment. W represents the current weight when entity N_i is input. $I(N_i)$ is the feature vector of N_i, and the values of $P(A_j|N_i)$ are used for this vector. $O(N_i)$ is the activation value of the output node as the response to $I(N_i)$. In the actual simulation, the number of the vector elements was 20. The feature words that are strongly related to the categories, including both positive and negative premise words, are selected. This selection is based on the assumption that only properties relevant to the context are used for induction (e.g., Shafto, Kemp, Baraff, Coley, and Tenenbaum, 2005).

Multilayer Neural Network Model: It is well known that this type of perceptron model cannot solve complex problems, such as linearly inseparable problems. The similarity-based induction processing that is indicated from the experimental data would appear to be beyond the computational capacity of a perceptron-based model. Accordingly, this study modifies the previous perceptron model to create a multi-layer model. The structure of the multilayer model is shown in Figure 3, and involves the following formula:

$$o^k = \sigma\left(\sum_i W_i f_i^k\right), \quad (7)$$

$$f_i^k = \sigma\left(\sum_j w_{ij} x_j^k\right), \quad (8)$$

where o^k denotes the activation value of the output node when the pattern of the kth conclusion N_k^c is input, W_i indicates the weights between the ith middle layer node and the output node, f_i^k represents the activation value of the ith middle layer node, and x_j^k denotes the jth element of the kth input pattern corresponding to $P(A_j|N_k^c)$. The activation strength of the output node o^k represents the likelihood of the conclusion. An ordinary sigmoid function was adopted as the activate function, σ, while the usual back propagation method was employed as the learning rule.

The premises and the conclusions were used for the learning process. In the learning process, the weight parameters are tuned so that the activation value of the output node o^k equals 1 in the case of positive premises, equals 0 in the case of negative premises, and equals each value obtained from the conclusion ratings of psychological experiment in the case of conclusions. The number of input nodes is 20, which is the same for the prior perceptron model. The number of middle layer nodes is set at 2, to keep the model as simple as possible.

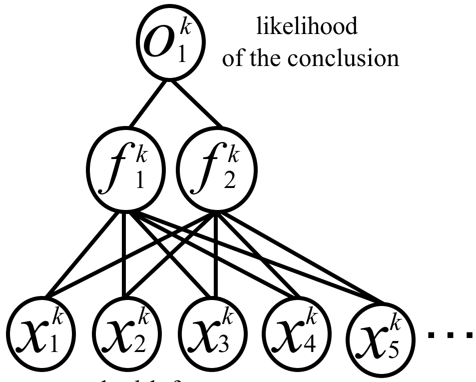

Figure 3. Structure of Multilayer Neural Network Model

Regression Model Based on Similarity Distance: This study also proposes another model of complex induction based on feature similarities, which is an extension of the Osherson et al. (1991) similarity regression model. In that model, the likelihood of a certain conclusion is computed using the linear summation of two kinds of similarities: the similarities that exist between the conclusion entity and the positive premises and the similarities that exist between the conclusion entity and the negative premises. These similarities are based on the features. The regression model proposed in this study has greater flexibility than the previous perceptron model, and is, therefore, also capable of simulating human performance for the complex task of induction based on feature similarities.

In that model, the likelihood of a conclusion including entity c_i, denoted as $v(c_i)$, is represented as follows:

$$v\left(N_i^c\right) = a\,\mathrm{SIM}_+\left(N_i^c\right) + b\,\mathrm{SIM}_-\left(N_i^c\right) + const \ , (9)$$

where $$\mathrm{SIM}_+\left(N_i^c\right) = \sum_j^{n+} e^{-\beta d_{ij}^+} \ , (10)$$

$$\mathrm{SIM}_-\left(N_i^c\right) = \sum_j^{n-} e^{-\beta d_{ij}^-} \ , (11)$$

$$d_{ij}^+ = \sum_k^m \left(P\left(A_k \mid N_i^c\right) - P\left(A_k \mid N_j^+\right)\right)^2 , (12)$$

$$d_{ij}^- = \sum_k^m \left(P\left(A_k \mid N_i^c\right) - P\left(A_k \mid N_j^-\right)\right)^2 . (13)$$

where a, b, and $const$ are parameters estimated from the likelihood of the positive premises (defined as value 7), the likelihood of the negative premises (defined as value 1), and the likelihood of each conclusion (value obtained from the experiment). N_j^+ is the entity of the positive premise, and N_j^- is the entity of the negative premise. $\mathrm{SIM}_+(N_i^c)$ and $\mathrm{SIM}_-(N_i^c)$ are the original functions for the feature similarities in this model. $\mathrm{SIM}_+(N_i^c)$ represents the similarities between the conclusion entity N_i^c and the positive premise entities, while $\mathrm{SIM}_-(N_i^c)$ denotes the similarities between N_i^c and the negative premise entities. ß is the only parameter in these functions. d_{ij}^+ and d_{ij}^- are also the original functions for word distance based on the feature words (denoted as a_k). Here, the number of feature words m is fixed to 20, matching the other models in this study. Although another similarity function was used in Osherson et al's model, as it required knowledge about some taxonomical categories and about the feature strengths of entities based on human ratings, that function would not allow the extended model to handle vast quantities of features that change dynamically according to context.

Evaluating the Model Simulations according to the Experimental Data

Simulations for all three models were executed. Table 4 shows correlation coefficients between the simulation results and the results from the psychological experiment of the within category condition, and F ratio (the fitness indices for the models). On the other hand, all correlation coefficients in the cases of the between category condition were larger than 0.7 and significant at $p < 0.01$, and all F ratio were also significant at $p < 0.05$. Considering these results, it is clear that the two models proposed in this study correlate well with both conditions, while the previous model only correlates with the between-category condition. These results indicate that the previous model is not able to simulate the experimental results obtained when the entities in positive and negative premises both belong to the same category.

Table 4. Correlation coefficients of the within category condition

	set1	set2	set3
Regression Model			
correlation coefficient	**0.939	**0.841	**0.936
F ratio	**30.06	**9.71	**28.58
Multilayer Model			
correlation coefficient	**0.968	**0.816	**0.899
F ratio	**135.94	**17.93	**37.93
Perceptron Model			
correlation coefficient	0.185n.s.	-0.09n.s.	0.36n.s.
F ratio	0.32n.s.	0.08n.s.	1.43n.s.

Discussion

The experiment results reported in this study are consistent with the hypothesis that people can estimate the likelihood of a conclusion, even when the entities in both positive and the negative premises belong to the same category, based on comparisons of the similarities between entities in positive premises and conclusions, and between those in negative premises and conclusions. Thus, these results provide verification of this hypothesis. The previous perceptron model, proposed by Sakamoto et al. (2005), was not able to simulate this experimental result.

From the comparisons of the simulation and experimental results, it is clear that the multilayer neural network model and the regression model based on similarity distance both correlated well with the results from the experiments, and that the performance of these models was better than the perceptron model. These results indicate the computational capacity of a perceptron model is not sufficient to handle cases where the induction of the appropriate category is difficult. On the other hand, the fact that the two proposed models both correlated well with the experimental data would seem to imply that two quite different approaches can both provide equally adequate accounts of the cognitive mechanisms underlying inductive reasoning. Despite their different theoretical underpinnings, however, the two proposed models would be represented in essentially identical ways in terms of support vector machines (SVMs) (Vapnik, 1995). SVMs are a kind of multilayer neural networks that provide solutions to the types of problems associated with multilayer neural networks, such as determining the number of multilayer-nodes and local minimum convergence. In order to avoid such problems, SVMs map feature patterns onto another dimensional space, where they become linearly partitioned. However, the computation of such complex mapping can also be achieved by a nonlinear-function computation, known as the kernel function. The kernel function is unconstrained except in instances where the function satisfies the mathematical condition of 'positive definiteness'. Returning to consider the regression model based on similarity distance proposed in this study, formulas (10) and (11) would correspond to the nonlinear mapping of the feature patterns $P\left(A_k \mid N_i^c\right)$ and $P\left(A_k \mid N_j\right)$. Moreover, as d_{ij} is a symmetric function of the feature pattern, $P\left(A_k \mid N_i^c\right)$ and $P\left(A_k \mid N_j\right)$, then $e^{-\beta d_{ij}}$ is also symmetric, that means that $e^{-\beta d_{ij}}$ has 'positive definiteness' and thus satisfies the condition of the kernel function. This instance indicates that the regression model can be represented as a SVM, that is, a multi-layer neural network. Consequently, the two models proposed in this study both have properties that are mathematically equivalent to the extent that after mapping feature patterns, which are nonlinear, linear partitions, as expressed in formula (7) and (8), can be achieved in the case of the regression model. Thus, despite their surface differences, the two models proposed in this study would both appear to be tapping into the basic cognitive mechanisms underlying the complex nature of inductive reasoning involving feature comparisons. In conclusion, in certain circumstances, people are able to estimate the likelihood of an argument's conclusion through complex processing involving comparisons of feature similarities between the entities in positive premises and the conclusion, and between the entities in negative premises and the conclusion.

Acknowledgments

This study was supported by the Tokyo Institute of Technology 21COE Program, "Framework for Systematization and Application of Large-scale Knowledge Resources".

The authors would like to thank Dr. T. Joyce, a post-doc with the 21COE-LKR, for his critical reading of manuscripts and valuable comments on an earlier draft.

References

Bruner, J. (1959). Inhelder and Piaget's The growth of logical thinking. *British Journal of Psychology, 50*, 363-370.

Buehner, M. J., & Cheng, P. W. (2005) Causal Learning. In. Holyoak, K. j., & Morrison, R. G. (Eds.) *The Cambridge Handbook of Thinking and Reasoning.* New York NY: Cambridge university press

Kameya, Y., & Sato, T. (2005). Computation of probabilistic relationship between concepts and their attributes using a statistical analysis of Japanese corpora. *Proceedings of Symposium on Large-scale Knowledge Resources: LKR2005*

Kudoh, T., & Matsumoto, Y. (2002). Japanese Dependency Analysis using Cascaded Chunking. *Proceedings of the 6th Conference on Natural Language Learning: CoNLL 2002*. 63-39.

Hanson, H. M., (1956). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *58,* 321-334.

Hofmann, T. (1999). Probabilistic latent semantic indexing. *Proceedings of the 22nd International Conference on Research and Development in Information Retrieval :SIGIR '99.* 50-57.

Osherson, D. N., Smith, E. E., Wilkie, O. Lopez, A., and Shafir, E. (1990). Category-Based Induction. *Psychological Review*, *97*, 2, 185-200.

Osherson, D. N., Stern, J., Wilkie, O., Stob, M., and Smith, E. E. (1991) Default Probability. *Cognitive Science, 15,* 251-269.

Pereira, F., Tishby, N., and Lee, L. (1993). Distributional clustering of English words. *Proceedings of the 31st Meeting of the Association for Computational Linguistics*. 183-190.

Sakamoto, K., Terai, A., and Nakagawa, M. (2005) Computational Models of Inductive Reasoning and Their Psychological Examination: Towards an Induction-Based Search-Engine. *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society.*

Shafto, P., Kemp, C., Baraff, L., Coley, J., and Tenenbaum, J. B. (2005). Context-sensitive induction. *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society.*

Sloman, A. T. (1993). Feature-Based Induction. *Cognitive Psychology, 25*, 231-280.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer.