Analysis of Cultural Differences in Pictogram Interpretations

Heeryon Cho

Abstract

The goal of this thesis is to present one concrete example of how a system can interpret and handle cultural differences in human-provided interpretations. In doing so, this thesis utilizes pictograms, a well-known iconic medium of communication which contains both the implicit cultural interpretations as well as explicit semantic interpretations. Cross-cultural pictogram interpretations are collected in two countries using a web survey, and the characteristics of pictogram interpretations are first analyzed. Based on the analyses, we identify three issues that need to be tackled to handle cultural differences in human-provided interpretations. The three issues are transformed into the following three actionable tasks:

1. A method of handling semantic ambiguity in pictogram interpretations is devised.

   Participants at both end of the communication channel must share common pictogram interpretation to communicate. However, because pictogram interpretation can be ambiguous, conveying intended meaning with a pictogram sometimes can be difficult. To assist human task of selecting pictograms more likely to be interpreted as intended, we propose a *semantic relevance measure* which calculates how relevant a pictogram is to a given interpretation. The proposed measure calculates the similarity measurement of two words in a pictogram set and the probability of each interpretation word in a given pictogram using pictogram interpretations and ratios gathered from a web survey. Then the proposed measure is applied to a pictogram retrieval

system. Moreover, the proposed measure is applied to categorized and weighted pictogram interpretation data to enhance pictogram retrieval performance. Five pictogram categories are defined based on the five first-level categories defined in the Concept Dictionary of the EDR Electronic Dictionary. Retrieval performance among not-categorized interpretations, categorized and not-weighted interpretations, and categorized and weighted interpretations using semantic relevance measure were compared, and the categorized and weighted semantic relevance approach showed better performance across a wider threshold interval.

2. How humans detect cultural differences in pictogram interpretations is investigated.

   Findings on how humans detect cultural differences in cross-cultural pictogram interpretations are reported. An open-answer web survey is conducted in the United States and Japan to collect U.S.-Japan pictogram interpretations. Thirty U.S.–Japan pictogram interpretations were used as stimuli for human cultural difference detection study. Three U.S. subjects and three Japanese subjects participated in the study to assess the degree of cultural differences in the thirty pictogram interpretations given in the questionnaire. Post-questionnaire interviews were conducted to elucidate the reasons behind the human cultural difference detection. The following factors were considered when humans detect cultural differences in cross-cultural pictogram interpretations: (i) similar or dissimilar interpretations in the two countries, (ii) percentage or ranking of the interpretations, (iii) conformity or variance of semantics within one country's interpretations, (iv) presence of proper nouns (e.g. country names), and (v) positive or negative connotation in the interpretations.

3. A method of handling cultural ambiguity in pictogram interpretations is devised.

   Pictograms have pictorial similarities with some object, and one who

can recognize the object depicted in the pictogram can interpret the meaning associated with the object. Some pictorial symbols, however, are interpreted differently across different cultures. To assist pictogram communication between intercultural participants, a method of automatically detecting cultural differences using two cultures' pictogram interpretation words and probabilities is proposed. Based on the human experiment which clarifies human cultural difference detection criteria, the human cultural difference detection criteria are formalized to define three detection inequalities which determine the cultural differences in two cultures' pictogram interpretations. The three inequalities are then merged using conjunctions to define a unified cultural difference detection function.

Nowadays, tags are prevalent form of metadata that are being used in various applications to describe, summarize, or impart additional meaning to the content to better assist content management by both humans and machines. Assuming that the two countries' pictogram interpretations are provided in words, assigning humans to detect cultural differences in two countries' pictogram interpretations is possible but not easy since it requires the human detector to have linguistic and cultural knowledge of both cultures. Hence, machine detection of cultural differences may be useful and helpful. With regard to the gathering of cross-cultural data, wide spread usage of tag-based applications today allows easy gathering of human interpretations in the form of tags. Moreover, if the user profiles and/or IP address information can be utilized to categorize tags, cross-cultural interpretation data can be realistically obtained via a networked system. Thus, the contributions of this work could be used as a basis for the construction of culturally-situated agent which can detect cultural differences in human-provided cross-cultural interpretations.

Acknowledgments

The road I have trodden has been a long and winding road, sometimes covered with thorn bushes, sometimes leading to a dead-end. This road, a tough, rough, rocky road, is called the "doctoral course," a notorious course known to test the traveler his/her perseverance, tenacity and flexibility (at the same time!), and courage and faith. My journey over this road was a difficult one, perhaps the most challenging one I have journeyed so far, but I managed to reach the goal thanks to the guidance of many wonderful people.

I am deeply indebted to my supervisor and my mentor, Professor Toru Ishida, for his continuous guidance, valuable advices, fruitful discussions, and heartwarming encouragements. We had very little common ground in the beginning, (I was from a different country, Korea, and different academic background, social science), and I suppose this made it more difficult for him to guide me; but he never gave up, and I am grateful for this. His stern and even relentless attitude toward research was exemplary, and I learned from him that maintaining a high standard in research was important and worth while. Other PhD students who have trodden before me have mentioned the word "role model" in their acknowledgements, and I must join them in addressing him as my role model. He is strict but kind, cool but passionate; those and more I wish to emulate. Thank you, Ishida-sensei, and I will do my best to someday reach you.

I am grateful to other members of my thesis committee at Kyoto University, Professor Kazuyuki Moriya and Professor Sadao Kurohashi, for the productive discussions and insightful research advices. I would also like to thank Professor Masatoshi Yoshikawa for his comments on the initial stage of my PhD research, and Assistant Professor Satoshi Oyama for the useful advice on the evaluation of the semantic relevance measure.

My PhD research took the form of joint research with NPO Pangaea, a nonprofit organization which aims to bridge children of the world using pictograms. I would like to express my sincere gratitude to Yumiko Mori and Toshiyuki Takasaki, president and vice president of NPO Pangaea, for giving me the opportunity to jointly conduct this exciting research.

Dr. Naomi Yamashita at NTT Communication Science Laboratories guided me through the human cultural difference detection experiment; without her expertise and sharp insight, the human experiment would not have been possible. Associate Professor Tomoko Koda at Osaka Institute of Technology gave helpful advices on cross-cultural research and encouraged me during my hardest times in PhD studies. Dr. Rieko Inaba at the Language Grid Project, National Institute of Information and Communications Technology (NICT), Japan, helped out with the initial research. I am grateful to them all.

I would also like to thank the current and former members of Ishida & Matsubara Laboratory for making my doctoral days memorable: I thank both the current faculty members, Associate Professor Shigeo Matsubara and Assistant Professor Hiromitsu Hattori, and the former faculty members, Associate Professor Hirofumi Yamaki, now at Nagoya University, and Associate Professor Hideyuki Nakanishi, now at Osaka University, for their guidance; I am grateful to Ms. Yoko Kubota and Ms. Terumi Kosugi, the coordinators of the Laboratory, for giving me both administrative and emotional support; many thanks to my colleagues, or "comrades" to be exact, Masahiro Tanaka, Donghui Lin, and Yuu Nakajima, for always being there; we toiled and cried and laughed together through our doctoral days, and those days I will cherish forever; I must thank Dr. Saeko Nomura, now at Cornell University, for supporting me when I was helpless like a puppy just after arriving in Kyoto; I also deeply appreciate Dr. Kosuke Nakatsuka, now at Sumitomo Electric, for introducing me to the exciting world of programming during his doctoral days at Ishida Laboratory; he will always be my "Master" and I, his "Grasshopper".

My family has been the prime source of my strength and energy, always encouraging me, invigorating me, and standing by me throughout my PhD journey: my father, Kyutae Cho, told me to “go with the flow”; my mother, Youngja Kim, reminded me to have a little courage and humor; my older sister, Yangjean Cho, cooked me the most delicious curry whenever I went back to Korea and taught me how to put things into perspective. I dedicate this thesis to my dearest family.

Last but not least, I thank God for lighting my way.

My PhD study was financially supported in full by the Japanese Government (Monbukagakusho) Scholarship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Objective

Infusion of information technology into the society has enabled people from all corners of the world to connect to each other virtually, enlarging opportunities to collaborate across borders. With the network infrastructure for cross-border collaboration set in place, researches to bridge multilingual, cross-cultural human participants are actively being carried out: [Ishida 06b] proposes a framework which seeks to overcome the human language barrier using language resources such as machine translators and electronic dictionaries; [Uchimoto 05] introduces a way to enable machine translator to assess its translation quality by proposing a similarity measure which has strong correlation with the quality of the translation result; [Yamashita 06] presents a method of predicting misconceptions present in multilingual computer-mediated communication by focusing on the discrepancies found in the semantic and syntactic thread of a multilingual electronic BBS log; [Koda 06] presents findings on a large scale online experiment which investigates how avatar's facial expressions are interpreted differently across cultures.

While the researches mentioned above seek ways to support computer-mediated intercultural collaboration, a long line of researches in the area of psychology and anthropology have shed light on cultural universals, i.e., factors universal to all cultures: [Osgood 57, Osgood 75] revealed a framework (Evaluation-Potency-Activity) which all humans share when differentiating affective meanings of signs; [Romney 97] proposed a method for visualizing semantic structure of human's internal cognitive representation and showed that English and Japanese share similar semantic structure of emotion terms; [Ekman 03] unmasked the seven emotions universally expressed by all cultures, which are anger, fear, disgust, surprise, sadness, happiness and contempt. Meanwhile, other researches have focused on identifying cultural differences exhibited across cultures: [Hofstede 05] formulated a cultural dimensions framework which identifies how power distance, collectivism vs. individualism, femininity vs. masculinity, and uncertainty avoidance are exhibited differently across cultures; [Hall 76] theorized the high-context vs. low-context aspect of communication in different cultures; [Nisbett 01] identified the cognitive differences in reasoning by East-Asians and Westerners through experimentation (the former was holistic while the latter was analytical).

On the other hand, in the area of computer-supported collaborative work (CSCW), [Setlock 04] showed that experimental groups with homogeneous and heterogeneous cultural backgrounds have different perceptions of the study task. In the area of human-computer interaction, [Olson 03] stated that culture influences both the process and the product of brainstorming, decision support systems, video as well as audio conferencing. Both works elucidated the cultural differences in computer-mediated interactions. In the area of artificial intelligence, [Chaudron 98] proposed a formal approach of cooperation which can be executed by software agents with different cultural backgrounds. With multifarious researches on culture unfolding in diverse academic fields, [Ishida 06a] envisioned a culturally-sensitive agent which can communicate culture.

This work continues in the line of research which deals with human cultural factor by presenting new findings on how humans perceive cultural differences, this time, in pictogram interpretations. Moreover, this work goes further to present a method of automatically detecting cultural differences in human-provided cross-cultural pictogram interpretations.

This work differs with the existing works on culture in the following aspects: it differs with [Ishida 06b, Uchimoto 05, Yamashita 06] in that the focus is on system-level handling of cultural differences rather than system-level handling of linguistic or semantic understandings (although linguistic and semantic handling must precede prior to the handling of cultural differences); it differs with [Koda 06] in that the interpretation data are gathered using free-answer questions rather than a predetermined set of (facial) interpretations, and that the interpretation data are used as stimuli to clarify human cultural difference detection criteria; it differs with existing psychological and anthropological findings on cultural universals or cultural differences in that the focus is on system-level handling of cultural factor rather than the discovery of fundamental theories of human nature (although the fundamental theories on culture might be used to explain the cultural phenomenon observed over the system); it differs with [Setlock 04, Olson 03] in that it deals with concrete objects, i.e., pictograms, rather than tasks, processes or interactions; and it differs with [Chaudron 98] in that a cultural difference detection methodology is proposed rather than a general framework for resolving conflicts arising from cultural differences.

To summarize, this work focuses on system-level handling and detection of cultural differences in human-provided cross-cultural pictogram interpretations. The findings in this research could be used to build a culturally-situated agent which can automatically detect cultural differences in cross-cultural pictogram interpretations; such agent can be viewed as one realization of culturally-sensitive agent envisioned in [Ishida 06a].

1.2 Approach and Issues

We use pictograms used in a real-world email system [Takasaki 07] to collect cross-cultural pictogram interpretations; the collected pictogram interpretations are then used as stimuli for elucidating the human cultural differences detection criteria. When collecting pictogram interpretations, we borrow the *semiosis* framework (the process in which something functions as a sign) proposed by Charles Morris. The semiosis process has commonly been regarded as involving three (or four) factors: that which acts as a sign, that which the sign refers to, and that effect on some interpreter in virtue of which the thing in question is a sign to that interpreter. These three components in semiosis may be called, respectively, the *sign vehicle*, the *designatum*, and the *interpretant*; the *interpreter* may be included as a fourth factor [Morris 38]. Figure 1.1 shows the semiosis of a pictogram (inside the box): a pictogram of a tower is the *sign vehicle*; the real-world towers shown in the photographs are the *designata*; the two humans standing in front of the pictogram are the *interpreters*; and the pictogram interpretations such as "Tokyo Tower (東京タワー)" and "Eiffel Tower" are the *interpretants*.

In order to build a culturally-situated agent which can automatically detect cultural differences in cross-cultural pictogram interpretations, this work follows the three steps described in Figure 1.1: first, cross-cultural pic-

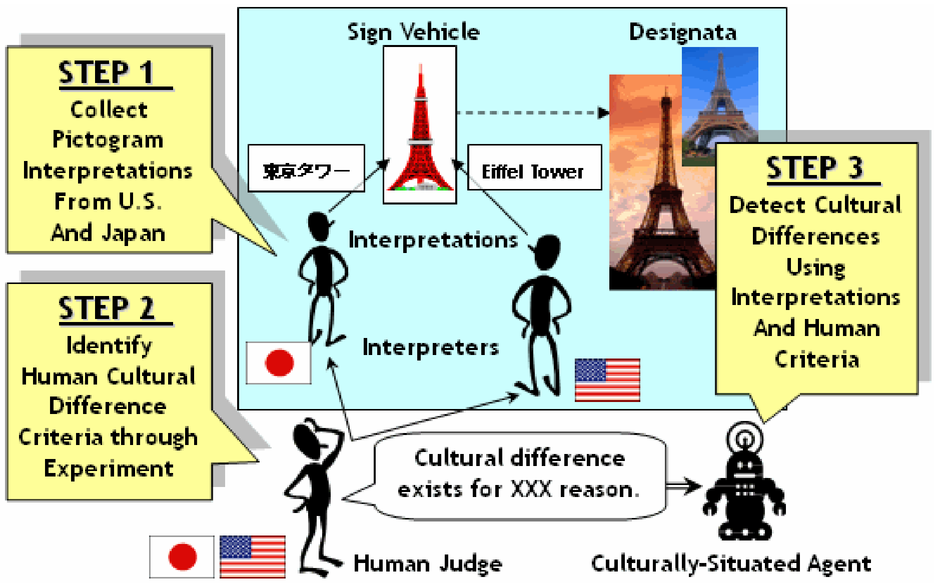


Figure 1.1: Semiosis of a pictogram (inside the box) and steps for building culturally-situated agent which can detect cultural differences in cross-cultural pictogram interpretations

togram interpretations are collected from the two countries, U.S. and Japan, which are known to have cultural differences [Hall 76, Hofstede 05]; then, human subject experiments are conducted to identify human cultural difference detection criteria using the cross-cultural pictogram interpretations as stimuli; finally, the human cultural difference detection criteria are formalized as mathematical functions to automatically detect cultural differences (the cross-cultural pictogram interpretations are used as data for calculating the values of the mathematical functions). Note that STEP 1 includes two research steps which must be cleared before proceeding to STEP 3: the first research step is analyzing the characteristics of pictogram interpretations (through this step, the semantic ambiguity and cultural ambiguity becomes evident); the second research step is devising a method to handle semantic ambiguity in pictogram interpretations.

Two major issues will be dealt with in this work: *semantic ambiguity* in one country's pictogram interpretations and *cultural ambiguity* in two countries' pictogram interpretations. This work will first tackle the semantic ambiguity in one country's pictogram interpretations and then extend the solution to cover cultural ambiguity in two countries' cross-cultural pictogram interpretations. In the process, this work will clarify the relationship between the two ambiguities.

1.3 Structure of the Thesis

The structure of this thesis is aligned with the steps shown in Figure 1.1 (STEP 1 is divided into two research steps). This thesis consists of six chapters with this chapter being the introductory chapter (Chapter 1). Note that Japanese words are presented in Japanese where necessary.

Chapter 2 analyzes the characteristics of pictogram interpretations; semantic ambiguity and cultural ambiguity in pictogram interpretations are identified as a result. An online survey is conducted in the United States and Japan to collect free-answer cross-cultural pictogram interpretations for 120 pictograms. The collected U.S.–Japan interpretation words are tallied to generate word comparison tables that list each country's interpretation words and ratios. The constitution of pictogram interpretation words of each pictogram is analyzed, first, with focus on one country, and next, with focus on two countries. Through these analyses, semantic ambiguity and cultural ambiguity in pictogram interpretations are identified.

Chapter 3 tackles the semantic ambiguity in pictogram interpretations using the *semantic relevance measure*. The semantic ambiguity (observed in one country's pictogram interpretations) includes two ambiguous interpretation characteristics: the first is one-to-many correspondence between pictogram-to-meaning, i.e., one pictogram has multiple interpretations (*polysemous interpretations*); the second is one-to-many correspondence between meaning-to-pictograms, i.e., multiple pictograms share common interpretation(s) (*shared interpretations*). The *semantic relevance measure* calculates how relevant a query is to a set of interpretation words of one pictogram; this enables the ranking of pictograms according to the query relevancy. This chapter also describes the categorization of interpretation words to enhance the pictogram retrieval performance.

Chapter 4 presents the findings of human cultural difference detection experiment. The experiment clarifies what criteria humans use when assessing the cultural differences in cross-cultural pictogram interpretations. Preliminary experiments were conducted to select 30 out of 120 pictograms with greater cultural differences. Six human subjects, three U.S. nationality and three Japanese, participated in a cultural difference detection experiment consisting of answering a questionnaire and answering to a post-questionnaire interview. Five human cultural difference detection criteria are identified as a result.

Chapter 5 formalizes the three out of five human cultural difference detection criteria to define three mathematical inequalities which calculate cultural differences in pictogram interpretations. The three inequalities are then merged using conjunctions to define a unified cultural difference detection function. The detection performance of the unified function is also evaluated. Finally, chapter 6 concludes this thesis.

Chapter 2

Semantic and Cultural Ambiguity in Pictogram Interpretations

This chapter describes the characteristics in pictogram interpretations. Cross-cultural pictogram interpretations are collected using a web survey and the characteristics in mono-cultural and cross-cultural pictogram interpretations are analyzed. Semantic ambiguity and cultural ambiguity in pictogram interpretations are discussed in detail using examples.

2.1 Introduction

Hand drawn images have long been used to convey messages, and are still being used as an effective iconic medium of representation. For instance, prehistoric drawings inside the Altamira Cave* tell us what wild animals lived during the ice age. Walking or standing human figures on the surface of a pedestrian traffic light alert us when to proceed or to stop. Hand drawn images are in essence iconic representation carrying semantic interpretation. We will call such images *pictograms* throughout this work. One of the most familiar pictograms used nowadays are universal signs such as road signs, direction boards at the airports, and symbols of sports played in the

*http://whc.unesco.org/en/list/310

Olympics. These pictograms are intended to convey particular information to a wide range of audiences. Because pictograms have clear pictorial similarities with some object [Marcus 03], people who can recognize the object depicted in the pictogram can recall meaning associated with the object. For this reason, many hopes were pinned to pictorial representations as a means to break the language barrier, and many attempts at universal picture writing system have been made and continues to this day.

Earlier in the 20th century, there were several attempts to develop universal languages based upon icons. Working in the 1920s and 1930s, Otto Neurath developed 'Isotype' which was strongly influenced by ideas from logical positivism and, particularly, the picture theory of meaning [Beardon 95]. Much effort has been put on developing pictograms for AAC (Augmentative and Alternative Communication). AAC assists people with severe communication disabilities to be more socially active in interpersonal interaction, education, employment, care management, and community activities. Sign language and Braille are good examples of AAC. Blissymbolics [Bliss 65] and PIC [Maharaj 80] are examples of pictogram communication systems used in AAC.

Meanwhile, yet another universal sign language and iconic communication systems were introduced: LoCoS was introduced by a graphic designer and sign designer [Vanhauer 07] and a mobile version of LoCoS was also proposed [Marcus 07]; CAILS, an experimental Computer Assisted Iconic Language System, dealt with three specific areas in communication (cross-linguistic, visual/spatial concept representation, and visual educational technique) [Champoux 00]; Minspeak [Baker 82] and IconText [Beardon 95] used a fixed set of icons and system-defined sentence generation procedures to create pictogram messages.

In this work, we look at a new kind of pictogram communication system, one that exchanges pictogram email messages via a network system [Mori 07, Takasaki 06, Takasaki 07]. The pictogram email system user involved in pictogram email exchange creates a pictogram message by selecting and combining one or more pictograms which are registered to the system. The pictogram email system provides a two-dimensional canvas interface where the user can freely place one or more pictograms onto the canvas. No system defined pictogram sentence generation rule is imposed on the user. Note that the registered pictograms are created by art major students who are novices at pictogram design; hence, the pictograms do not guarantee a fixed, unique interpretation. Moreover, new pictograms are continuously added to the system resulting in the registered pictogram repository to grow.

With this in mind, we will look at how these pictograms are interpreted and what characteristics these pictogram interpretations have by analyzing the pictogram interpretations gathered from a web survey.

2.2 Pictogram Web Survey

Objective

An online pictogram survey was conducted to clarify how pictograms are interpreted by people living in one country and identify what characteristics, if any, those pictogram interpretations have. Moreover, the online survey was extended to gather interpretations from participants living in two countries in order to understand whether differences in pictogram interpretations exist in two countries.

Method

A pictogram survey, which asks the meaning of 120 pictograms used in the email system [Takasaki 07], was conducted to respondents in the U.S. and Japan via the WWW from October 1, 2005 to November 30, 2006.† Human respondents were shown a webpage similar to Figure 2.1 which contains 10 pictograms per page, and were asked to write the meaning of each pictogram inside the text field provided below the pictogram. Each time a set of 10 pictograms was shown at random, and the respondents could choose and answer as many question sets they liked. The maximum question sets

†The URL of the online survey is `http://www.pangaean.org/iconsurvey/`.

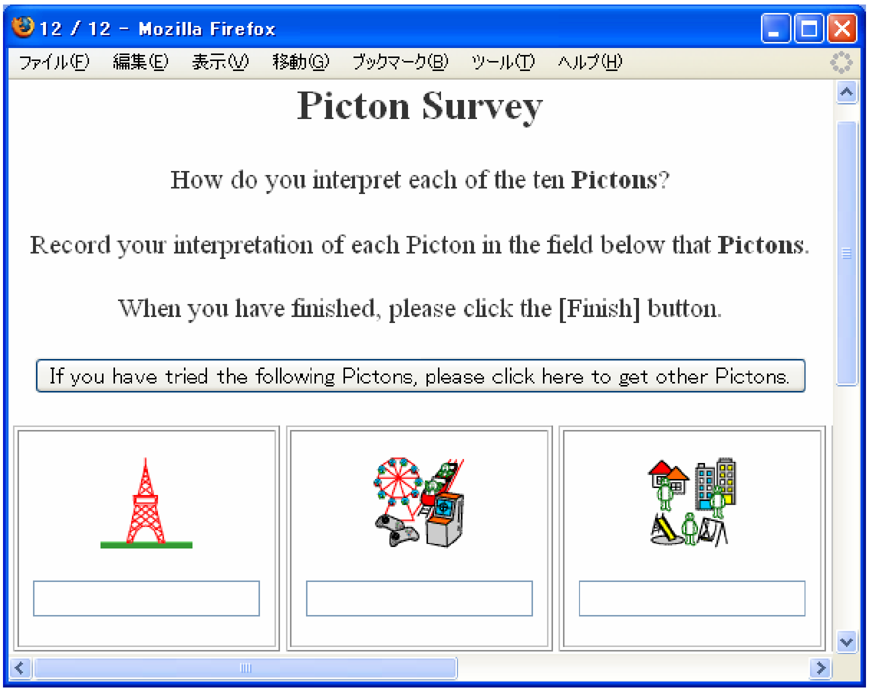


Figure 2.1: A screenshot of the online pictogram survey page (Three out of ten pictograms are shown)

a respondent could answer were 12 sets which contain a total of 120 pictograms.

Data

A total of 543 respondents in Japan and 935 respondents in the U.S. participated in the survey. An average of 97 interpretations consisting of Japanese words or phrases (duplicate expressions included) and an average of 147 interpretations consisting of English words or phrases (duplicate expressions included) were collected for each pictogram. For each pictogram, unique interpretation words or phrases were listed for each language, and the occurrences of those unique words were tallied to calculate the total frequency of each word. An example of U.S.–Japan word count result for one of the pictogram surveyed is shown in Table 2.1. The left two columns show interpretation words and frequencies collected from the U.S. respondents. The right two columns show interpretation words and frequencies collected from the Japanese respondents.

For example, U.S. interpretation word “dancing” placed at the top has a frequency of “51”. This means that fifty-one U.S. respondents wrote “dancing” as the meaning of the pictogram displayed at the top of the table. Comparative charts that were created for analyses contain the original Japanese words as they are, but in this paper we translate all Japanese words into English for readability. A Japanese-English dictionary, EDICT[‡], was used for translation.

Words and phrases which were not listed in the dictionary (including colloquial expressions) were translated by humans. Parentheses following each English translation of the Japanese word in Table 2.1 contain the original Japanese word expressed in alphabet (in *italics*) and the Japanese character construction of the original term: “hr” denotes *hiragana*, “kt” denotes *katakana*, and “kj” denotes *kanji*. Italicized Japanese term and its character construction are delimited by a colon(:).

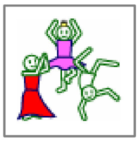
2.3 Semantic Ambiguity in Pictogram Interpretations

The analysis of U.S. pictogram interpretation words revealed two characteristics evident in pictogram interpretation. Firstly, all 120 pictograms had more than one pictogram interpretation making them polysemous. That is, each pictogram had more than one meaning to its image. Secondly, some pictograms shared common interpretation(s) with one another. That is, some pictograms shared exactly the same interpretation word(s) with one another.

Here we take up eight pictograms to show the above mentioned characteristics in more detail. For the first characteristic, we will call it *polysemous pictogram interpretation*. For the second, we will call it *shared pictogram*

[‡] The URL is `http://www.csse.monash.edu.au/~jwb/j_edict.html`.

Table 2.1: An example of U.S.–Japan interpretation words and frequencies



U.S.	FREQ.	JAPAN	FREQ.
dancing	51	dance (*dansu*: kt)	45
dance	25	dance (*odori*: kj+hr)	13
gymnastics	7	dance (*odori*: hr)	6
dancers	6	dance (*dansu*: hr)	2
ballet	5	fun (*tanoshii*: kj+hr)	2
play	5	dance (*odoru*: hr)	1
cheerleaders	4	circus (*sa–kasu*: kt)	1
danceing	3	dancer (*dansa–*: kt)	1
playing	3	performance (*pafo–mansu*: kt)	1
family	2	clown (*piero*: kt)	1
friends	2	theatrical play (*engeki*: kj)	1
acrobatics	1	hobby (*shumi*: kj)	1
ballerina show	1	battle (*tatakai*: kj+hr)	1
cheerleading	1	gymnastics (*taisou*: kj)	1
cherrleaders	1	dance (*odoru*: kj+hr)	1
dance class	1	dance (*odori*: kj+hr), dance (*dansu*: kt)	1
dancing triplets	1	everyone getting along well (*minnanakayoku*: hr+kj+hr)	1
exercise	1	rhythmic sports gymnastics (*shintaisou*: kj)	1
flexable	1		
girls playing	1		
hurting eachother	1		
i like to dance	1		
play time	1		
playin	1		
TOTAL FREQUENCY	126	TOTAL FREQUENCY	81

interpretation. To guide our explanation, we categorize the interpretation words into the following seven categories: (i) People, (ii) Place, (iii) Time, (iv) State, (v) Action, (vi) Object, and (vii) Abstract category. The images

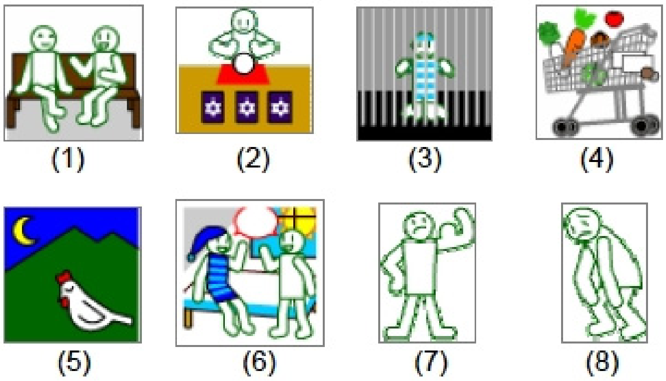


Figure 2.2: Pictograms having polysemous interpretations (See Table 2.2 for interpretations)

of the pictograms are shown in Figure 2.2. The interpretations of Figure 2.2 pictograms are organized in Table 2.2. The interpretation words shared by more than one pictogram are marked in *italics* in both the body text and the table.

People. Pictograms containing human figures (Figure 2.2 (1), (2), (3), (6), (7), (8)) can have interpretations explaining something about a person or a group of people. Interpretation words like "*friends*, fortune teller, magician, prisoner, criminal, strong man, bodybuilder, tired person" all explain specific kind of person or group of people.

Place. Interpretations may focus on the setting or background of the pictogram rather than the object occupying the center of the setting. Figure 2.2 (1), (3), (4), (7) contain human figure(s) or an object like a shopping cart in the center, but rather than focusing on these central objects, words like "church, jail, prison, grocery store, market, gym" all denote specific place or setting related to the central objects.

Time. Concept of time can be perceived through the pictogram and interpreted. Figure 2.2 (5), (6) have interpretations like "night, *morning*, dawn, evening, bed time, day and night" which all convey specific moment of the day.

Table 2.2: Polysemous interpretations and shared interpretations (marked in *italics*) found in Figure 2.2 pictograms and their interpretation categories

Pic.	Interpretation	Category
(1)	*friends* / church / *happy*, *talking* / talk, *play*	Person / Place / State / Action
(2)	fortune teller, magician / fortune telling, magic	Person / Abstract
(3)	prisoner, criminal / jail, prison / stuck, raining	Person / Place / State
(4)	grocery store, market / basket full, *healthy* / food, cart, vegetables / shopping	Place / State / Object / Abstract
(5)	night, *morning*, dawn, evening, bed time / sleeping / sleep, *wake up* / chicken, moon	Time / State / Action / Object
(6)	*friends* / *morning*, day and night / *happy*, *talking* / *play*, *wake up*	Person / Time / State / Action
(7)	strong man, bodybuilder / gym / strong, *healthy*, *hurt* / exercise / muscle / strength	Person / Place / State / Action / Object / Abstract
(8)	tired person / tired, weak, *hurt*	Person / State

State. States of some objects (including humans) are interpreted and described. Figure 2.2 (1), (3), (4), (5), (6), (7), (8) contain interpretations like "*happy*, *talking*, stuck, raining, basket full, *healthy*, sleeping, strong, *hurt*, tired, weak" which all convey some state of the given object.

Action. Words explaining actions of the human figure or some animal are included as interpretations. Figure 2.2 (1), (5), (6), (7) include interpretations like "talk, *play*, sleep, *wake up*, exercise" which all signify some form of action.

Object. Physical objects depicted in the pictogram are noticed and indicated. Figure 2.2 (4), (5), (7) include interpretations like "food, cart, vegetables, chicken, moon, muscle," and they all point to some physical object(s) depicted in the pictograms.

Abstract. Finally, objects depicted in the pictogram may suggest more abstract concept. Figure 2.2 (2), (4), (7) include interpretations like "fortune telling, magic, shopping, strength" which are the result of object-to-concept association. Crystal ball and cards signify fortune telling or magic, shopping cart signifies shopping, and muscle signifies strength.

We showed the two characteristics of pictogram interpretation, *polysemous pictogram interpretation* and *shared pictogram interpretation*, by presenting the actual interpretation words exhibiting those characteristics as examples. We believe such varied interpretations are due to differences in how each respondent places his or her focus of attention to each pictogram.

2.4 Cultural Ambiguity in Pictogram Interpretations

As mentioned in section 2.2, an online pictogram survey was conducted in the U.S. and Japan to understand how different cultures interpret pictograms. The selection of the two countries is based on the fact that chances of finding cultural differences in pictogram interpretation would be higher if we choose cultures that have greater cultural differences. Since existing literatures on cross-cultural studies have found the two countries' cultures to be distinct in many aspects[Hall 76, Hofstede 05, Vatrapu 07], we proceed with our survey in the two countries.

Tables comparing English and Japanese pictogram interpretation words and frequencies (similar to Table 2.1, but containing the original Japanese words and phrases) were created for each of the 120 surveyed pictograms. To determine whether culture-specific interpretations were present, three human judges independently analyzed the 120 English–Japanese pictogram interpretations for cultural differences. Two judges were Japanese and one judge was Korean. All three judges had college level Japanese and English proficiency. After reviewing the 120 pictogram interpretation words, each of the three judges found 8, 21(Korean judge), and 28 pictograms to have culturally different interpretations. Nineteen pictograms were found to have culturally different interpretations by two or more judges. Seven pictograms were found to have culturally different interpretations by all three judges.

We give details of the nineteen pictograms which were judged by two or more judges to have culturally different interpretations by the U.S. respondents and Japanese respondents. To guide our explanation, we divide the pictograms into the following groups: (i) Gesture, (ii) Gender and Color, (iii) Time, (iv) Space, (v) Familiar Scenery, and (vi) Facial Expression. The top five frequent U.S.–Japan pictogram interpretation words are listed for each pictogram along with their percentages. The percentage (PCT, %) of each word or phrase is calculated by dividing the interpretation word frequency with the total frequency. For example, the percentage of the word "dancing" in Table 2.1 can be calculated as $(51/126) * 100 = 40.48\%$. For all Japanese interpretation words, English translations, alphabetical expressions of the Japanese terms (in *italics*), and Japanese character constructions are provided as those shown in Table 2.1.

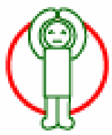
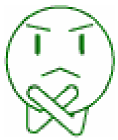
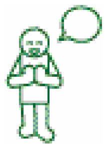
(i) Gesture

A pictogram of a person holding up one's hands above one's head to form a circle-like shape (Table 2.3 top) was interpreted as "exercise, jump rope, exercising, yoga, dance, stretch" by a majority of U.S. respondents whereas a majority of Japanese respondents interpreted it as "OK, circle, correct, all right, bingo". The U.S. interpretations center on exercise-related concept while the Japanese interpretations center on agreement-related concept.

Likewise, pictogram of a person crossing one's arms to form an "X" mark (Table 2.3 middle) was interpreted as "mad, angry, anger, frustrated, upset" by a majority of U.S. respondents whereas a majority of Japanese respondents interpreted it as "no good, wrong[§], no, miss, don't". The U.S. interpretations revolve around the concept which deals with negative emotions while the Japanese interpretations revolve around the concept which

[§]Although EDICT lists three entries to the Japanese term "*batsu*," English translation fitting the context of the pictogram (the "X" gesture) could not be found so a more appropriate human translation is given.

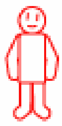
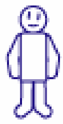
Table 2.3: Example of cultural ambiguity in gesture

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	exercise	19.47	OK (originally in alphabet)	18.81
	jump rope	5.31	circle (*maru*: hr)	9.90
	happy	4.43	correct (*seikai*: kj)	8.91
	exercising	3.54	O.K. (*okke–*: kt)	7.92
	yoga	3.54	O.K. (*iiyo*: hr)	5.94
	mad	31.90	no good (*dame*: hr)	26.73
	angry	30.17	no good (*dame*: kt)	11.88
	no	4.31	wrong* (*batsu*: hr)	5.94
	stubborn	3.45	wrong* (*batsu*: kt)	3.96
	anger	1.72	no (*iie*: hr)	2.97
	talking	10.19	thank you (*arigatou*: hr)	6.33
	praying	9.55	please (*onegai*: hr+kj+hr)	6.33
	thinking	8.28	to speak (*hanasu*: kj+hr)	5.06
	speaking	5.10	soliloquy (*hitorigoto*: hr)	3.80
	lonely	3.19	soliloquy (*hitorigoto*: kj+hr+kj)	3.80

deals with prohibition or criticism.

As for the pictogram that shows a standing person placing hands together while a speech balloon hangs next to the head (Table 2.3 bottom), approximately 40% of both the U.S. and Japanese respondents interpreted it as some kind of a speech act (“talking, speaking” and “to speak, soliloquy” respectively). At the same time, however, 14.6% of U.S. respondents interpreted it as “praying, pray, prayer” while 17.7% of Japanese respondents interpreted it as “thank you, please”. These differences in the interpretations of the three pictograms, we think, are due to the differences in how gestures are interpreted in the U.S. and Japan. The body gestures expressing a circle or a cross are gestures well-recognized in Japan which respectively indicate that something is correct or wrong. However, such gesture is not recognized in the United States: therefore we suppose that the circle depicted in the pictogram was perceived as an expression of motion while the “X” was perceived as crossing of one’s arms (hence, the stubborn or angry gesture) by the U.S. respondents. The important point to note is that while the two countries’ overall interpretations of the top and middle pictogram dif-

Table 2.4: Example of cultural ambiguity in gender and color

PICTOGRAM	U.S.	PCT(%)	JAPAN	PCT(%)
	woman	29.05	woman (*onnanohito*: kj+hr+kj)	28.00
	man	11.49	woman (*josei*: kj)	27.00
	mom	8.78	woman (*onna*: kj)	10.00
	dad	7.43	mother (*okaasan*: hr)	5.00
	adult	5.41	mother (*okaasan*: hr+kj+hr)	3.00
	man	34.23	man (*otokonohito*: kj+hr+kj)	27.72
	dad	10.07	male (*dansei*: kj)	26.73
	woman	8.05	man (*otoko*: kj)	9.90
	adult	6.04	father (*otousan*: hr)	3.96
	mom	5.37	father (*otousan*: hr+kj+hr)	3.96

Note: Top pictogram is drawn in red line whereas bottom pictogram is drawn in blue line.

fer greatly, the bottom pictogram contains a mixture of both the differing interpretations (“praying” vs. “thank you”) and the common interpretation shared by the two countries (“talking” and “to speak”).

(ii) Gender and Color

“The color red denotes women and the color blue denotes men” is a prevalent notion in Japan, but it is not so in the U.S. as indicated by the pictogram interpretations. While 92% of the Japanese respondents interpreted the red human figure (Table 2.4 top) as “woman, mother, adult female, sister, girl” which all contain the female gender concept, 31.8% of the U.S. respondents interpreted it as “man, dad, father, boy, male” which all contain the male gender concept. Strong agreement in interpretation was reached by the Japanese respondents, but not by the U.S. respondents. The remaining 8% of the Japanese interpretations consisted of “adult” and “person” which lacks any gender concept, and “boy” that was answered by one Japanese respondent. As for the remaining U.S. respondents, most of them interpreted the red human figure similarly as the majority of Japanese did as some kind of a female person. Small portion of the U.S. respondents interpreted it as “adult, person, teenager, grown up, parent”.

Likewise, the blue human figure (Table 2.4 bottom) was interpreted by

93% of the Japanese respondents as “man, male, father, adult male, brother, boy” which all contain the male gender concept. In contrast, 20.8% of the U.S. respondents interpreted it as some person with the female gender, i.e., “woman, mom, big girl, female, old women”. Only one Japanese respondent interpreted it as a “girl”. The remaining U.S. respondents interpreted the blue human figure similarly as the Japanese as some kind of a person with the male gender. In sum, it can be concluded that the correlation of color and gender (red denotes female and blue denotes male) is evident in the Japanese interpretations, but not in the U.S. interpretations.

The important point to notice is that while the Japanese interpretations center on a single gender concept, i.e. the concept of male or female, the U.S. interpretations include both gender concepts for each pictogram leading to a greater ambiguity in interpretation.

(iii) Time

In both countries, pictograms containing clock image(s) (Table 2.5) were interpreted as some kind of a concept relating to time, but the first ranking interpretations were different between the two countries.

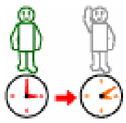
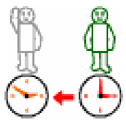
Starting with the top pictogram in Table 2.5, the first ranking U.S. interpretation was “late (11.38%)” whereas the first raking Japanese interpretation was “the future (16.83%).” Since the third ranking Japanese interpretation shown in the table is also “the future (5.94%),” it can be combined with the first ranking interpretation to yield a total percentage of 22.77%. Similar interpretations to the first ranking U.S. interpretation (“late”) also existed below the ranking, which include “5 min. late, late or later, you are late” which were each answered by one U.S. respondent. A total of 4.19% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were “future, forward in time, past to future”. A total of 5% of Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were “lateness (*chikoku*: hr, kj), to be late (*okureru*: kj+hr)”. Common interpretations shared by the two countries included “10 minutes later, time passes”.

As for the middle pictogram in Table 2.5, the first ranking U.S. and Japanese interpretations were “on time (16.87%)” and “now (11.11%)” respectively. Since the second and fourth ranking Japanese interpretations shown in Table 2.5 are also “now (10.00% and 4.44%),” they can be combined to yield a total percentage of 25.55%. Similar interpretations to the first ranking U.S. interpretation (“on time”) existed below the ranking, which include “be on time, you are on time”. A total of 6.63% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were “now, present, current time, present time”. A total of 5.56% Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were “just (*choudo*: hr), on time (*ontaimu*: kt, *jikandoori*: kj+kt).” Common interpretation shared by the two countries was “time”.

For the bottom pictogram in Table 2.5, the first ranking U.S. and Japanese interpretations were “early (12.27%)” and “the past (18.09%)” respectively. Since the third ranking Japanese interpretation shown in the table is also “the past (7.45%),” it can be combined to yield a total percentage of 25.54%. Similar interpretations to the first ranking U.S. interpretation (“early”) existed below the ranking which include “10 minutes early, 5 min early, early or earlier, someone’s early, you are early”. A total of 4.91% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were “past, backward in time, future to past”. A total of 2.11% Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were “arrived early (*hayakutsuichatta*: kj+hr+kj+hr), arrived 10 minutes ago (*juppunmaenikimashita*: num+kj+hr).” Common interpretation shared by the two countries was “10 minutes ago”.

In sum, the three pictograms containing clock image(s) were interpreted by the U.S. respondents as “late, on time, early” whereas the Japanese respondents interpreted them as “future, present, past”. It can be said that the U.S. interpretations deal with a concept of appointment in relation to time while the Japanese interpretations deal with temporal relations along the time axis. The important point to notice is that the basic time concept

Table 2.5: Example of cultural ambiguity in time

PICTOGRAM	U.S.	PCT(%)	JAPAN	PCT(%)
	late	11.38	the future (*mirai*: kj)	16.83
	time	10.18	10 minutes later (*juppungo*: num+kj)	9.90
	10 minutes	3.59	the future (*mirai*: hr)	5.94
	later	2.99	afterwards (*atode*: hr)	3.96
	future	2.40	time passes (*jikangasusumu*: kj+hr+kj+hr)	2.97
	on time	16.87	now (*genzai*: kj)	11.11
	time	12.65	now (*ima*: kj)	10.00
	now	3.61	time (*jikan*: kj)	6.67
	what time is it	3.61	now (*ima*: hr)	4.44
	clock	3.01	time (*jikan*: hr)	4.44
	early	12.27	the past (*kako*: kj)	18.09
	before	4.29	10 minutes ago (*juppunmae*: num+kj)	11.70
	past	3.68	the past (*kako*: hr)	7.45
	time	3.68	time is turned back (*jikangamodoru*: kj+hr+kj+hr)	4.26
	late	3.07	some time ago (*sakki*: hr)	3.19

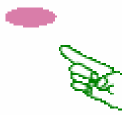
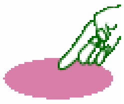
is shared by the two countries, but the detailed interpretations that unfold around the time concept differs as manifested by the two countries' first ranking interpretations.

(iv) Space

Two pictograms portraying an index finger pointing to a specific place were interpreted differently by the two countries. Although both countries' interpretations revolved around the concept of space, the perspectives held by the respondents were different. We focus on the first ranking interpretations (as we did in the prior time related pictograms) to highlight the differences.

For the Table 2.6 top pictogram, 18.88% of the U.S. respondents interpreted the finger's direction to be pointing "up" whereas 30.63% of the

Table 2.6: Example of cultural ambiguity in space

PICTOGRAM	U.S.	PCT(%)	JAPAN	PCT(%)
	up	18.88	there (*asoko*: hr)	30.63
	there	14.69	that (*are*: hr)	27.03
	far	6.29	there (*acchi*: hr)	9.91
	over there	4.90	above (*ue*: kj)	5.41
	point	4.90	far (*tooi*: kj+hr)	3.60
	down	18.88	here (*koko*: hr)	36.04
	here	16.08	this (*kore*: hr)	26.13
	near	6.99	this direction (*kocchi*: hr)	5.41
	big	3.50	below (*shita*: kj)	5.41
	low	3.50	near (*chikai*: kj+hr)	3.60

Japanese respondents interpreted it as pointing to “there”. Since the third ranking Japanese interpretation shown in the table is also “there (9.91%),” it can be combined with the first ranking Japanese interpretation to yield a total percentage of 40.54%. Similar interpretations to the first ranking U.S. interpretation (“up”) existed below the ranking which add up to 11.89%: example interpretations include “high, pointing up, up/above, above, look up, up high”. Combining the first ranking U.S. interpretations with the similar, below ranking interpretations, the percentage of the major U.S. interpretation “up” adds up to 30.77%.

For the Table 2.6 bottom pictogram, 18.88% of the U.S. respondents interpreted it as “down” whereas 36.04% of the Japanese respondents interpreted it as “here”. Since the third ranking Japanese interpretation, “this direction (5.41%),” shown in the table contains similar meaning to the first ranking Japanese interpretation, it can be combined to yield a total percentage of 41.45%. Similar interpretations to the first ranking U.S. interpretation (“down”) existed below the ranking which add up to 20.98%: example interpretations include “low, pointing down, down/below, below, look down, down low”. Combining the first ranking U.S. interpretations with the similar, below ranking interpretations, the percentage of the major U.S. interpretation “down” adds up to 39.86%.

Major interpretation observed in one country was also observed in the other country, but with a lower percentage. The first ranking Japanese interpretations "there" and "here" for the top and bottom pictogram (Table 2.6) were also observed within the U.S. interpretations (totaled 24.48% and 25.87% respectively): example interpretations include "there, over there, go there, look there, spot there" and "here, right here, come here, look here, spot here". On the other hand, Japanese interpretations similar to the first ranking U.S. interpretations "up" and "down" (top and bottom pictogram in Table 2.6) were totaled 9% and 8.11% respectively: example interpretations include "above (*ue*: kj, hr), high (*takai*: kj+hr)" and "below (*shita*: kj, hr), low (*hikui*: kj+hr)." Common interpretations shared by the two countries were "far" and "near" respectively for the top and bottom pictogram in Table 2.6.

In sum, the two pictograms depicting a finger pointing to a certain direction were interpreted as "up, down" by the U.S. respondents whereas the Japanese respondents interpreted them as "there, here". It can be said that the U.S. interpretations contain a vertical perspective of space while the Japanese interpretations contain a horizontal perspective of space. The important point to notice is that while the basic concept of space is shared by the two countries, the major (or the first ranking) interpretations vary as evidenced by "up vs. there" and "down vs. here".

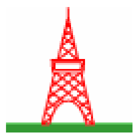
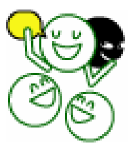
(v) Familiar Scenery

In some cases, the U.S. and Japanese respondents recalled familiar scenes from the visual scenery depicted in the pictograms. These recalled scenes varied according to culture.

In the case of the top pictogram in Table 2.7, nearly half (43.08%)[¶] of the U.S. respondents interpreted the red tower as the "Eiffel Tower" while nearly half (47.83%) of the Japanese respondents interpreted it as the "Tokyo Tower". Apparently, the respondents recalled specific instances of the tower they were familiar with. None of the U.S. respondents submit-

[¶]Twelve misspelled versions of the "Eiffel" were observed in the U.S. interpretations including the fourth ranking "Eifel Tower".

Table 2.7: Example of cultural ambiguity in familiar scenery

PICTOGRAM	U.S.	PCT(%)	JAPAN	PCT(%)
	Eiffel Tower	19.23	Tokyo Tower (*toukyoutawa–*: kj+kt)	44.57
	Paris	19.23	tower (*tawa–*: kt)	23.91
	tower	15.38	tower (*tou*: kj)	7.61
	Eifel Tower	4.62	Eiffel Tower (*efferutou*: kt+kj)	6.52
	France	4.62	tower (*denpatou*: kj)	3.26
	winner	30.63	athletic meet (*undoukai*: kj)	36.59
	winning	6.88	number one (*ichiban*: kj)	8.54
	champion	5.63	overall victory (*yuushou*: kj)	6.88
	first place	5.00	number one (*ichiban*: hr)	3.66
	cheering	3.13	first place prize (*ittoushou*: kj)	3.66
	friends	9.38	liar (*usotsuki*: hr)	7.89
	party	8.13	to tell a lie (*usowotsuku*: hr)	5.26
	gossip	3.75	lie (*uso*: kj)	3.95
	happy	3.13	lie (*uso*: hr)	2.63
	happy group	3.13	malicious gossip (*kageguchi*: hr)	2.63

ted “Tokyo Tower” as the interpretation, but 7.6% of the Japanese respondents submitted “Eiffel Tower” as the interpretation. Common interpretation shared by the two countries was “tower”.

In the case of the middle pictogram in Table 2.7, the first ranking interpretations given by the U.S. and Japanese respondents were “winner (30.63%)” and “athletic meet (36.59%)” respectively. Note that such athletic meet depicted in the pictogram is a regularly held school event in Japan. Therefore, it is reasonable to assume that the Japanese respondents associated the pictogram’s visual scenery to the school hosted athletic meet. Similar interpretations shared by the two countries (U.S. / Japan) were “winning / overall victory” and “champion, first place / number one”.

The case with Table 2.7 bottom pictogram should be given greater attention since the two countries’ interpretations vary greatly, almost going the opposite direction. While most of the U.S. respondents interpreted the pictogram to mean “friends, party, happy, happy group, laughing, having fun”, etc. which all indicate a cheery, positive scene, most of the Japanese respondents interpreted it to mean “liar, to tell a lie, lie, malicious gossip, split personality, vicious, to deceive, scheming”, etc. which all indicate a shady, negative image. We assume that the Japanese respondents have interpreted the black face on the upper right corner as a person having a malicious intent or an ulterior motive. Hence, the negative interpretation is given.

In contrast, we assume that the U.S. respondents interpreted the black face to be an African American, and as a result, interpreted the four faces as a group of people with varying ethnic background. Since people from diverse ethnic groups are portrayed as chatting together, it is a desirable scene, and thus positive interpretations are derived. Such interpretation, however, may be difficult to come out from Japanese respondents, since Japan is an ethnically homogeneous country, and almost all people (excluding the foreigners) belong to the same ethnic group. Therefore, it is more natural to interpret a different face color as signifying the person’s state of mind. The important point to mention with regard to the three pictograms dealing with familiar scenery is that they contain a mixture of different interpretation patterns: while the top and middle pictogram respectively contains a common underlying concept such as “tower” and “winning”, the bottom pictogram contains vastly varying interpretations.

(vi) Facial Expression

Facial expressions were interpreted differently not only between the two countries, but also among the respondents within the same country. Starting with the top pictogram in Table 2.8, the greatest common U.S. interpretation was “whistling, whistle (25.16%)” whereas the greatest common Japanese interpretation was “feigning ignorance, pretending not to know (30.38%).” Other varying interpretations were given by the members of each country. For instance, U.S. respondents interpreted as “curious, kiss, relieved, sad, startled, embarrassed, snobby” while Japanese respondents interpreted as “to pout, to deceive, boring, to jeer, to get angry, to tell a lie, to bluff”.

Table 2.8: Example of cultural ambiguity in facial expression

PICTOGRAM	U.S.	PCT(%)	JAPAN	PCT(%)
	whistling	13.21	feigning ignorance (*shiranpuri*: hr)	5.06
	whistle	10.06	to be peevish (*suneru*: hr)	5.06
	no	5.66	hmm (*hun*: hr)	5.06
	annoyed	2.52	turning suddenly (*pui*: hr)	5.06
	ignore	2.52	whistle (*kuchibue*: kj)	5.06
	scared	18.01	cold (*samui*: kj+hr)	27.18
	cold	10.56	cold (*samui*: hr)	23.30
	worried	10.56	scary (*kowai*: hr)	9.71
	nervous	9.94	scary (*kowai*: kj+hr)	4.85
	sad	9.94	trembling (*buruburu*: hr)	2.91
	happy	6.49	good-looking (*kakkoii*: hr)	31.7
	mean	5.19	handsome (*hansamu*: kt)	8.65
	smart	4.55	boast (*jiman*: kj)	2.88
	boy	3.90	nice man (*iiotoko*: hr+kj)	1.92
	mischievous	3.90	ahem (*ehhen*: hr)	1.92
	happy	25.64	cute (*kawaii*: hr)	42.72
	girl	3.85	pretty (*kirei*: hr)	5.83
	nice	3.85	cute (*kawaii*: kj+hr)	2.91
	pretty	3.85	beautiful person (*bijin*: kj)	2.91
	sweet	3.85	chuckling (*ufufu*: hr)	1.94
	happy	8.05	pretty (*kirei*: hr)	16.49
	in love	4.70	beautiful person (*bijin*: kj)	13.40
	cute	4.03	cute (*kawaii*: hr)	8.25
	pretty	3.36	beautiful (*utsukushii*: kj+hr)	4.12
	sweet	3.36	a prim girl (*osumashi*: hr)	2.06
	sly	11.95	to make fun of (*bakanisuru*: hr)	3.00
	sneaky	11.32	bitter smile (*nigawarai*: kj+hr)	3.00
	happy	6.92	doubt (*utagai*: hr)	2.00
	cool	2.52	grinning (*niyaniya*: hr)	2.00
	shy	2.52	broadly grinning (*niyari*: hr)	2.00

As for the second pictogram in Table 2.8, the top two interpretations in the U.S. and Japan were “scared (24.84%), cold (11.8%)” and “cold (60.95%), scared (24.27%)” respectively. Notice that although both countries share the same two interpretations “cold, scared”, the first and second commonly shared interpretations are reversed between the two countries.

Moving to the third, fourth, and fifth pictogram in Table 2.8, the first ranking interpretations for each of the three pictograms were “happy, happy, happy” by the U.S. respondents, and “good-looking, cute, pretty” by the Japanese respondents. Japanese respondents tend to interpret the outer appearance of the face while U.S. respondents interpreted the state of the mind projected through the face. The fourth pictogram had, compared to the other two pictograms, a relatively high agreement in interpretation within each country with 25.64% answering “happy” in the U.S. and 42.72% answering “cute” in Japan. As for the remaining two pictograms (Table 2.8 third and fifth), low agreement on interpretation was reached especially among the U.S. respondents.

The last pictogram shown at the bottom of Table 2.8 consists of widely varying interpretations not only between the two countries, but also among the members within each country. It can mean “sly, sneaky, happy, cool, shy” in the U.S. while “to make fun of, bitter, doubt, grinning, broadly grinning” in Japan. However, most of the Japanese interpretations contained negative connotations whereas the U.S. interpretations contained both negative and positive connotations. For example, U.S. interpretations such as “pleased, clever, glad, proud, smart” were positive interpretations that never appeared in the Japanese interpretations.

In sum, pictograms containing facial expressions can have varying interpretations not only between the two countries, but also among the members of the same country. Note that although there were other pictograms that depicted facial expressions, for example, pictograms depicting a crying face or an angry face, these pictograms with negative facial expressions were interpreted similarly by the two countries. There were no cultural differences in the interpretations.

2.5 Summary

We looked at the details of culturally different interpretations in nineteen pictograms. Although each pictogram contained a specific cultural difference in interpretation, we think that these cultural differences can be categorized into the following three patterns. We give examples of each pattern.

- The basic concept captured by the two cultures are the same, but the perspectives on that concept are different.
  E.g. The concept related to time, space, tower, face are captured by both cultures (U.S. and Japan), but how they are perceived vary.
  - [Table 2.5] *late* vs. *future*, *on time* vs. *now*, *early* vs. *past*
  - [Table 2.6] *up* vs. *there*, *down* vs. *here*
  - [Table 2.7 top] *Eiffel Tower* vs. *Tokyo Tower*
  - [Table 2.8 third, fourth, fifth] *happy* vs. *good-looking, cute, pretty*

- The basic concept(s) are only partially shared by the two cultures.
  - [Table 2.3 bottom] *talking, to speak* is shared, but not *praying*, *thank you*
  - [Table 2.4] *woman* is shared, but not *man* and vice versa
  - [Table 2.7 middle] *winning, overall victory* is shared, but not *athletic meet*

- There is no common concept captured by the two cultures.
  E.g. A gesture is recognized by one culture, but not by the other.
  - [Table 2.3 top & middle] *exercise* vs. *O.K.*, *mad* vs. *no good*
  - [Table 2.8 top] *whistle* vs. *feigning ignorance*

  E.g. Specific environment leads to specific recognition.
  - [Table 2.7 bottom] *friends* vs. *liar*

As a first step to understanding how pictograms are interpreted in different cultures, a pictogram web survey asking the meaning of pictograms was conducted in the U.S. and Japan. Three human judges independently analyzed the survey results containing English–Japanese pictogram interpretations to see whether cultural differences in pictogram interpretation exist between the two countries. As a result, nineteen out of the 120 surveyed pictograms were judged by two or more human judges to have culturally different interpretations.

The analysis of the nineteen culturally different pictogram interpretations confirmed the following three patterns of cultural differences in pictogram interpretation:

1. Two cultures share the same underlying concept, but have different perspectives on the concept.
2. Two cultures only partially share the same underlying concept.
3. Two cultures do not share any common underlying concept.

These findings can be utilized in designing a pictogram retrieval system which can detect and notify the cultural differences in pictogram interpretation.

Chapter 3

Semantic Relevance Measure and Categorized Semantics

This chapter introduces a measure which calculates the semantic relevance between a word and a set of pictogram interpretation words. The semantic relevance measure defines similarity of two words in a pool of pictograms and the word similarities are weighted according to the probability of each word. The measure is applied to a pictogram retrieval system and its performance is evaluated. Moreover, a priori categorization of pictogram interpretation words is proposed to enhance retrieval performance.

3.1 Introduction

Tags are prevalent form of metadata used in various applications today, describing, summarizing, or imparting additional meaning to the content to better assist content management by both humans and machines. Among various applications that incorporate tags, we focus on a pictogram email system which allows children to communicate to one another using pictogram messages [Takasaki 06, Takasaki 07]. Our goal is to support smooth pictogram communication between children, and to realize this, we focus on the pictogram selection stage where children select individual pictograms to create pictogram messages.

Pictogram is an icon which has a clear pictorial similarity with some object [Marcus 03], and one who can recognize the object depicted in the pictogram can interpret the meaning associated with the object. Pictorial symbols, however, are not universally interpretable. A simple design like an arrow is often used to show direction, but there is no reason to believe that arrows suggest directionality to all people; they might also be taken as a symbol for war or bad luck [Kolers 69]. Since the selection of pictogram in the pictogram email system is done with the purpose of conveying certain meaning to the communicating counterpart, pictogram that was selected must carry intended meaning to both the sender and receiver of communication; that is, selected pictogram must be relevant to participants at both end of communication channel in order for the pictogram communication to be successful.

To assist pictogram selection, we propose *semantic relevance measure* which calculates how relevant a pictogram is to a given interpretation. We also propose a priori *categorization and weighting* of pictogram interpretations. Related research unifies the browsing by tags and visual features for intuitive exploration of image databases [Aurnhammer 06]. Our approach first categorizes and weights pictogram interpretations and then calculates *semantic relevance* to rank relevant pictograms in terms of given interpretation. We define *pictogram categories* used for categorizing pictogram interpretations, by appropriating the first level categories defined in the Concept Dictionary of the EDR Electronic Dictionary [EDR 02]. We show that categorized and weighted semantic relevance pictogram retrieval approach returns better result compared to not-categorized, not-weighted approaches.

3.2 Ambiguity in Pictogram Interpretation Revisited

A twenty-five month pictogram web survey was conducted from October 1st, 2005 to November 7th, 2007 to collect free-answer English pictogram interpretation words or phrases from respondents living in the United States.

Based on the unique username and IP address pair, a total of 1,576 respondents participated in the survey. Details of the earlier survey can be found in [Cho 07a, Cho 07b].

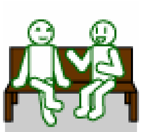
3.2.1 Polysemous Interpretation

To analyze the characteristics of pictogram interpretation, English interpretation words or phrases collected from the web survey were first tallied according to unique interpretation words, and then phrasal interpretations and misspellings were discarded. An example of tallied pictogram interpretations is shown in Table 3.1. As shown, a pictogram can have various interpretations which include both similar and different-meaning words. For example, words like *talking, talk, conversation, chatting, chat, communication* are all similar action-related words describing the act of speaking. Other action-related words are *date, flirt, sit, love, play*. When the focus shifts to the people depicted in the pictogram, however, the pictogram is interpreted as *friends* or *family*. Or it can be interpreted as some kind of place such as *church*, or as an emotional state such as *happy*. One way to organize mixed interpretations containing both similar and different meaning words is to group them into related categories. We use the Headconcept Dictionary and Concept Classification Dictionary in the Concept Dictionary of the EDR Electronic Dictionary [EDR 02] to categorize pictogram interpretation words*.

The EDR Electronic Dictionary was developed for advanced processing of natural language by computers, and is composed of five types of dictionaries (Word, Bilingual, Concept, Co-occurrence, and Technical Ter-

*SUMO ontology [Niles 01] was another candidate for categorizing pictogram interpretations, but we chose the EDR for three reasons: (1) we needed to handle both Japanese and English pictogram interpretations, and the EDR provides both English and Japanese headconcepts; (2) the first level classes directly located below the SUMO ontology's `Entity Class` are `Abstract Class` and `Physical Class`, and these classes, we think, are less human-friendly compared to the first level categories in the EDR; (3) EDR was specifically developed for natural language processing, therefore was more compatible to our research purpose which involves communication.

Table 3.1: An example of tallied pictogram interpretations

PICTOGRAM	INTERPRETATION	FREQ.	RATIO
	talking	57	0.388
	talk	28	0.190
	conversation	19	0.129
	friends	15	0.102
	chatting	13	0.088
	chat	2	0.014
	date	2	0.014
	flirt	2	0.014
	sit	2	0.014
	church	1	0.007
	communication	1	0.007
	family	1	0.007
	friend	1	0.007
	friendly	1	0.007
	love	1	0.007
	play	1	0.007
	TOTAL	147	1.002

minology), as well as the EDR Corpus. The Concept Dictionary contains information on the approximately 410,000 concepts listed in the Word Dictionary and is divided according to information type into the Headconcept Dictionary, the Concept Classification Dictionary, and the Concept Description Dictionary. The Headconcept Dictionary describes information on the concepts themselves. The Concept Classification Dictionary describes the super-sub relations among the approximately 410,000 concepts. The "super-sub" relation refers to the inclusion relation between concepts, and the set of interlinked concepts can be regarded as a type of thesaurus. The Concept Description Dictionary describes the semantic (binary) relations, such as 'agent', 'implement', and 'place', between concepts that co-occur in a sentence [EDR 02]. We define *five pictogram categories* using the five first level categories defined directly below the root concept of the Concept Dictionary. The five categories on the first level are:

(a) human or subject whose behavior (actions) resembles that of a human
(b) {matter} an affair
(c) event/occurrence
(d) location/locale/place
(e) time

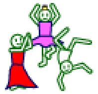
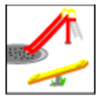
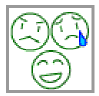
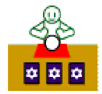
We define the five pictogram categories as (a) AGENT, (b) MATTER, (c) EVENT, (d) LOCATION, and (e) TIME; each maps to the aforementioned first level categories respectively. We now have five pictogram categories, and the pictogram interpretation words are categorized into the appropriate pictogram category through the following steps: first, concept identifier(s) are obtained by searching the English headconcept(s) in the Headconcept Dictionary that matches the interpretation word; then, using the concept identifier, the first level category or categories are obtained by climbing up the super-sub relations defined in the Concept Classification Dictionary. Note that since (i) there can be more than one concept identifier matching the headconcept, and (ii) for some concept identifier the Concept Classification Dictionary allows multiple inheritances, there can be more than one category that a single interpretation word can be categorized into.

Table 3.2 shows interpretation words of nine pictograms categorized into the five pictogram categories. Each column lists interpretation words of each pictogram, and each row displays interpretation words assigned to one of the five pictogram categories. Although some interpretation words are actually categorized into multiple pictogram categories, Table 3.2 only displays each word assigned to a single *major category*[†]. We now look at interpretations in each category.

(a) AGENT Pictograms containing human figures can trigger interpretations explaining something about a person or a group of people. Table 3.2 AGENT row contains words like *family, people, crowd, fortuneteller, gypsy,*

[†]One interpretation word may be repeatedly categorized to the same category. Major category is the most repeatedly obtained category. More explanation is given in section 3.3.2.

Table 3.2: Polysemous interpretations (each column) and shared interpretations (boldface type)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
									
AGENT					**family**	**family**		people **family** crowd	fortuneteller gypsy magician
MATTER	good morning **good night** **hello**	**moon** good evening **dancing** **good night** **hello**	**moon** **good night**	chicken **picture**		**dancing** ballet		**card** **picture**	**card** crystal ball magic
EVENT	**talking** **happy** **play**	**talking** **happy** **play**	**sleeping** dream peaceful	**sleeping** relaxed	**talking** date **play**	dance jumping **play**	slide fun **play**	crying **happy** mixed	bowling guess **play**
LOCATION							park	theater	
TIME	**morning** day **bedtime**	**night** **evening**	**night** **bedtime**	**morning** **night** **evening**					future

or *magician*, which all explain specific kind of person or people.

(b) MATTER Concrete objects or objective subjects are indicated. Table 3.2 MATTER row contains words like *good morning, good night, hello, moon, good evening, dancing, chicken, picture, ballet, card, crystal ball*, or *magic*, which point to some physical object(s) or subject depicted in the pictogram.

(c) EVENT Actions or states are illustrated. Table 3.2 EVENT row contains words like *talking, happy, play, sleeping, dream, peaceful, relaxed, date, dance, jumping, slide, fun, crying, mixed, bowling*, or *guess*, which all convey present states or ongoing actions.

(d) LOCATION Place, setting or background is highlighted rather than the object occupying the center or the foreground of the setting. Table 3.2 LOCATION row contains words like *park* or *theater*, which all indicate specific place or setting related to the central object(s).

(e) TIME Time-related concepts are perceived and captured. Table 3.2 TIME row contains words like *morning, day, bedtime, night, evening*, or *future*, which all describe specific moments in time.

Categorizing the words into five pictogram categories elucidates two key aspects of polysemy in pictogram interpretation. Firstly, interpretations that spread across different categories lead to different meanings. For example, interpretation words in Table 3.2 column (8) include AGENT category's *crowd* and *family* which describes a specific kind of people, EVENT category's *crying* and *happy* which describes action or emotional state, and LOCATION category's *theater* which describes a place for presenting and seeing a show; and they all mean different things.

Secondly, while interpretation words placed within the same category may contain related words such as *sleeping* and *dream* in Table 3.2 column (3) EVENT row, or *dance* and *jumping* in column (6) EVENT row, contrasting words sometimes coexist within the same category. For example, Table 3.2 column (4) TIME row contains both *morning* and *night* which are contrasting time-related words, and column (1) MATTER row contains both *good morning* and *good night* which are also contrasting greeting words. While the words in Table 3.2 column (2) TIME row are varied yet located closer to one another along the time axis (*night* and *evening*), the words in column (4) TIME row are placed at the opposite end (*morning* and *night*), leading to a stronger contrast in words in terms of time concepts. To summarize the first and second findings, it can be said that polysemy in pictogram interpretation is generally observed 'across' different categories, but sometimes strong contrast in interpretations can be observed 'within' the same category.

When a pictogram having polysemous interpretations is used in communication, there exists a chance that the sender and receiver might interpret the same pictogram differently. For instance, in the case of pictogram in Table 3.2 column (4), it could be interpreted differently as *morning* and *night* by the sender and receiver respectively. One way to assist the sender in the selection of a pictogram with higher chance of conveying the intended meaning is to display possible interpretations of the pictogram. If various possible interpretations are presented, the sender can speculate on the receiver's interpretation before selecting a pictogram. For example, if the sender knows a priori that Table 3.2 pictogram (4) can be interpreted as both *morning* and *night*, he or she can guess ahead that it might be interpreted differently by the receiver, and avoid choosing the pictogram. The displaying of possible pictogram interpretations is the first pictogram selection assistance we propose in tackling the issue of polysemy or one-to-many correspondence in pictogram and pictogram interpretations.

3.2.2 Shared Interpretation

A single pictogram may have various interpretations, but these interpretations are not necessarily exclusive to one pictogram; sometimes two or more pictograms may share the same interpretation(s). Interpretation words indicated in boldface type in Table 3.2 are such interpretations shared by more than one pictogram: for example, Table 3.2 (5), (6), and (8) share *family* (AGENT row); (1), (2), and (3) share *good night* (EVENT row); and (2), (3), and (4) share *night* (TIME row) and so on. The fact that multiple pictograms can share common interpretation implies that each one of those pictograms can be interpreted as such. The degree to which each is interpreted as the common interpretation, however, may vary according to the pictogram. For instance, Table 3.2 pictograms (1), (2), and (3) can all be interpreted as *good night* (MATTER row), but (1) can also be interpreted as *good morning* while (2) can also be interpreted as *good evening*. Furthermore, if we move down the table to the TIME row, we see that (1) contains *morning* as time-related concept while (2) and (3) contain *night*.

Suppose two people A and B each selects pictograms (1) and (2) respectively to send a “good night” message to person C. Upon receiving the message, C however, may interpret A’s message as “good morning” while interpret B’s message as “good night”. Even though A and B both intend to convey a “good night” message, it may not be the case that C will interpret both pictograms as such; this is because the degree of the shared interpretation may vary across different pictograms. Such degree difference, we think, is affected by both the probability of the shared interpretation of each pictogram and the remaining interpretations within each pictogram. For example, the probability of *good night* in Table 3.2 column (1) and column (2), and the remaining interpretations such as *good morning* in column (1) versus *good evening* in column (2), both affect how strong (1) and (2) pictogram can be interpreted as *good night*.

When two or more pictograms share the same interpretation I, the degree to which each pictogram may be interpreted as I may vary. If the degree of interpretation I for each pictogram is known, selecting the pictogram with the greatest degree of I will increase the chance of conveying I compared to other candidate pictograms. Hence, one way to assist pictogram selection among multiple pictograms sharing the same interpretation is to rank those pictograms according to the degree of relevancy to a given interpretation. Once the sender is presented with the ranked pictograms listed according to the relevancy of interpretation I, the sender can estimate which pictogram is most likely to be interpreted as I. In order to rank pictograms according to the relevancy of certain interpretation, some kind metric which measures the relevancy of a pictogram to a given interpretation is needed; to this end we propose the *semantic relevance measure*. The calculation of semantic relevance and the ranking of pictograms according to the interpretation relevancy is the second pictogram selection assistance we propose to tackle the issue of synonymy or many-to-one correspondence in pictograms and pictogram interpretation.

3.3 Categorical Semantic Relevance

Previous section identified ambiguities in pictogram interpretation, namely, polysemy and synonymy in pictogram interpretation, and issues involved in the usage of such ambiguous pictograms in communication. In this section, we first define the *semantic relevance measure* which calculates relevancy value of each pictogram when a pictogram interpretation is given, and then we propose categorization and weighting of the pictogram interpretation words prior to the semantic relevance calculation to improve pictogram retrieval performance.

3.3.1 Semantic Relevance Measure

We assume that pictograms each have a list of interpretation words and frequencies as the one given in Table 1. Each unique interpretation word has a frequency, and each word frequency indicates the number of people who answered the pictogram to have that interpretation. The ratio or the probability of an interpretation word, which can be calculated by dividing the word frequency by the total word frequency of that pictogram, indicates how much support people give to that interpretation. For example, in the case of Table 3.1 pictogram, it can be said that more people support *talking* (57 out of 147) as the interpretation for that pictogram than *sit* (2 out of 147). The higher the ratio or the probability of a specific interpretation word in the pictogram, the more that pictogram is accepted by people for that interpretation. We define *semantic relevance measure* of a pictogram to be the measure of relevancy between a word query and interpretation words of a pictogram.

Let $w_1, w_2, ..., w_n$ be interpretation words of pictogram $pict$. Let the probability of each interpretation word in a pictogram to be $P(w_1 \mid pict), P(w_2 \mid pict), ..., P(w_n \mid pict)$. For example, the probability of the interpretation word *talking* for Table 3.1 pictogram can be calculated as $P(talking \mid pict_{Table3.1}) = 54/104$. Then the simplest equation that assesses the relevancy of a pictogram $pict$ in relation to a word query w_i can be defined as follows:

$$P(w_i \mid pict) \tag{3.1}$$

This equation, however, does not take into account the similarity of interpretation words. For instance, when "talking" is given as query, pictograms having similar interpretation word like *gossiping*, but not *talking*, fail to be measured as relevant when only the probability is considered. To solve this, we need to define some kind of similarity, or $Similarity(w_i, w_j)$, between interpretation words. Using the similarity, we can define the *semantic relevance measure* or $SemanticRelevance(w_i, pict)$ as follows:

$$SemanticRelevance(w_i, pict) = \sum_j P(w_j \mid pict) Similarity(w_i, w_j) \tag{3.2}$$

There are several similarity measures. We draw upon the definition of similarity given by Lin [Lin 98] which states that similarity between A and B is measured by the ratio between the information needed to state the commonality of A and B and the information needed to fully describe what A and B are. Here, we calculate the similarity of w_i and w_j by counting how many pictograms contain certain interpretation words. When there is a pictogram set $Pict$ having an interpretation word w_i, the similarity between interpretation words w_i and w_j can be defined as follows. $|Pict(w_i) \cap Pict(w_j)|$ is the number of pictograms having both w_i and w_j as interpretation words. $|Pict(w_i) \cup Pict(w_j)|$ is the number of pictograms having either w_i or w_j as interpretation word.

$$Similarity(w_i, w_j) = \frac{|Pict(w_i) \cap Pict(w_j)|}{|Pict(w_i) \cup Pict(w_j)|} \tag{3.3}$$

Based on (2) and (3), the *semantic relevance* or the measure of relevancy to return pictogram $pict$ when word w_i is input as query can be calculated as follows:

$$SemanticRelevance(w_i, pict) = \sum_j P(w_j \mid pict) \frac{|Pict(w_i) \cap Pict(w_j)|}{|Pict(w_i) \cup Pict(w_j)|} \tag{3.4}$$

The calculated semantic relevance values fall between zero and one, which denotes that either a pictogram is completely irrelevant to the interpretation or completely relevant. By sorting the semantic relevance values, pictograms can be ranked from very relevant (value close to 1) to not so relevant (value close to 0). As the value nears zero, we observe that pictograms become less and less relevant; hence, a cutoff-point is needed to discard the less relevant pictograms. This is done to prevent the usage of pictograms which could be ineffective in conveying a given interpretation. Setting an ideal cutoff-point that satisfies all word query and pictogram interpretations, however, is difficult since for different queries, the ideal cutoff-point fluctuates.

One reason for the cutoff-point fluctuation is that all words contained in each pictogram, regardless of how great or small each interpretation word is related to the query, influence the semantic relevance calculation, and as a result, contribute to the final semantic relevance value. For example, let's say that we want to find a pictogram which can convey the meaning "friend" or "friends". Pictogram in Table 3.1 could be a candidate since it contains both *friends* and *friend* with a total ratio close to 0.11. When the semantic relevance value is calculated, however, the equation takes into account not only the two words, but all the remaining interpretation words including *talking*, *church*, *play*, and so forth. So, one way to remedy the dispersion of interpretation is to selectively use a set of interpretation words more related to the query, and use those selected words in the semantic relevance calculation to reduce the effect of less-related interpretation words influencing the calculation. With this prediction, we test the semantic relevance calculation on categorized interpretations.

3.3.2 Categorizing the Pictogram Interpretations

Pictogram interpretation words are categorized into the five pictogram categories described in section 3.2.1. Note that one interpretation word may be categorized into multiple pictogram categories since a word may link to multiple concept identifiers via the same headconcept or via multiple inheritances. For example, in the case of the word (headconcept) *park*, three kinds of pictogram categories are obtained repeatedly: LOCATION category six times, MATTER category five times, and EVENT category four times. In such case of multiple category acquisition, we use all categories since we cannot accurately guess on the single correct category intended by each respondent who participated in the web survey.

3.3.3 Weighting the Pictogram Interpretations

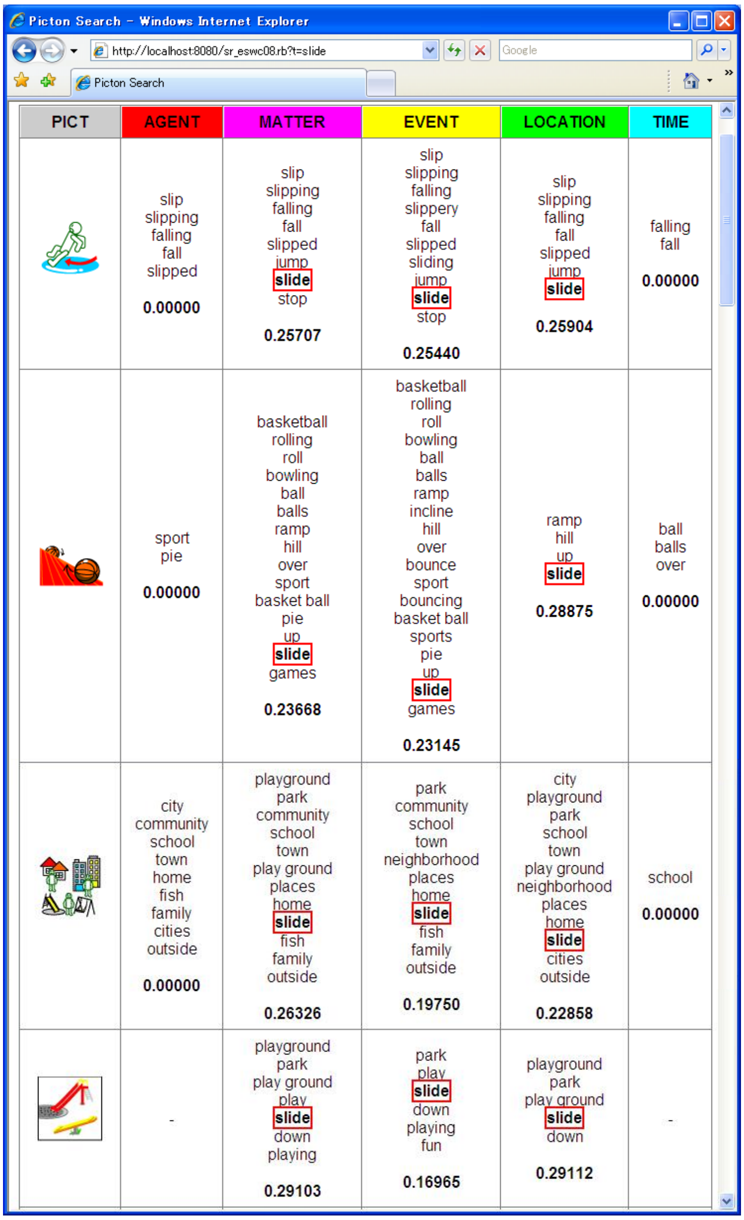
Although we cannot correctly decide on the single, intended category of a word, we can calculate the ratio of the pictogram category of each word. For example, in the case of the word *park*, the LOCATION category has the most number of repeated categories (six). Next is the MATTER category (five) followed by the EVENT category (four). We can utilize such category constitution by calculating the ratio of the repeated categories and assigning the ratio as weights to the word in a given category. For example, the word *park* can be assigned to LOCATION, MATTER and EVENT category, and for each category, weights of 6/15, 5/15 and 4/15 can be assigned to the word. Consequently, the *major category* of the interpretation word *park* will be LOCATION.

3.3.4 Ranking the Result

Applying the semantic relevance calculation to categorized interpretations will return five *categorical semantic relevance values* for each pictogram. We take the highest categorical semantic relevance value and compare it with the cutoff point to determine whether the pictogram is relevant or not. Once the relevant pictograms are selected, the selected pictograms are then sorted according to the semantic relevance value of the query's major category. For example, if the query is "park", then the relevant pictograms are first selected using the highest categorical semantic relevance value of each pictogram, and once the relevant pictograms are selected, the pictograms are ranked according to the categorical semantic relevance value of the query's major category, which in this case is the LOCATION category. The resulting list of pictograms is a ranked list of pictograms starting with the most relevant pictogram on top.

3.3.5 Prototype Implementation

We implemented a prototype web-based pictogram retrieval system which returns a list of relevant pictograms in the descending order of the query's major category's semantic relevance values when a word query is given as input. The categorized and weighted pictogram interpretation words for 120 pictograms were given to the system as data to calculate the categorical semantic relevance values. Figure 3.1 shows a list of retrieved pictograms for the query "slide." Note that the retrieved pictograms are sorted according to the EVENT category's semantic relevance values since the major category of the query "slide" is EVENT. When the pictograms are ranked according to the LOCATION category's semantic relevance values, however, the ranking changes with the fourth pictogram with the highest LOCATION value (0.29112) jumping to the top. This difference in the categorical semantic relevance value will be utilized in the pictogram sentence generator described in section 3.5.



PICT	AGENT	MATTER	EVENT	LOCATION	TIME
	slip slipping falling fall slipped **0.00000**	slip slipping falling fall slipped jump **slide** stop **0.25707**	slip slipping falling slippery fall slipped sliding jump **slide** stop **0.25440**	slip slipping falling fall slipped jump **slide** **0.25904**	falling fall **0.00000**
	sport pie **0.00000**	basketball rolling roll bowling ball balls ramp hill over sport basket ball pie up **slide** games **0.23668**	basketball rolling roll bowling ball balls ramp incline hill over bounce sport bouncing basket ball sports pie up **slide** games **0.23145**	ramp hill up **slide** **0.28875**	ball balls over **0.00000**
	city community school town home fish family cities outside **0.00000**	playground park community school town play ground places home **slide** fish family outside **0.26326**	park community school town neighborhood places home **slide** fish family outside **0.19750**	city playground park school town play ground neighborhood places home **slide** cities outside **0.22858**	school **0.00000**
	-	playground park play ground play **slide** down playing **0.29103**	park play **slide** down playing fun **0.16965**	playground park play ground **slide** down **0.29112**	-

Figure 3.1: A screenshot of a prototype web-based pictogram retrieval system which uses the categorized and weighted semantic relevance approach (Results for the query “slide” is displayed)

The prototype system implements the two design principles discussed in sections 3.2.1 and 3.2.2 which deals with one-to-many (i.e. polysemous) and many-to-one (i.e. shared) relationship between pictogram and pictogram interpretations. Evaluation of the proposed method is described next.

3.4 Evaluation

3.4.1 Comparison of the Four Approaches

Three pictogram retrieval approaches that singly uses or combines the semantic relevance measure, word categorization, and word weighting were evaluated. The baseline for comparison was a simple string match of the word query to the pictogram interpretation words with probabilities greater than the cutoff point. This is the same as selecting pictograms with $P(w_j|e) >$ *cutoff point* where w_j equals the query. A relevant pictogram set was constructed by five human judges, and retrieval tasks were performed using the four approaches: (1) baseline string match approach, (2) not-categorized semantic relevance approach, (3) categorized semantic relevance approach, and (4) categorized and weighted semantic relevance approach.

3.4.2 Relevant Pictogram Set Construction

Five human judges were employed in the construction of a relevant pictogram set which consists of 188 pictogram interpretation words and a ranked list of relevant pictograms for each word. The judges were all undergraduate students and they were paid for their tasks. The relevant pictogram set was constructed through the following steps:

[STEP 1] COLLECTING HUMAN ASSESSMENT DATA: A questionnaire containing 188 pictogram interpretation words[‡] with candidate pictograms[§], each listing all interpretation words (similar to the second column in Table 3.1) was given to the five human judges, and for each interpretation word, the human judges were asked to (i) judge whether each candidate pictogram could be interpreted as the given word (i.e. judged either as relevant or not relevant), and (ii) if judged as relevant, write down the ranking among the relevant pictograms.

[STEP 2] JUDGING AND RANKING RELEVANT PICTOGRAMS: The five judges' assessment data were averaged and variances were calculated to select and rank relevant pictograms for each interpretation word. If three or more people judged the pictogram to be relevant, the pictogram was selected as relevant. Otherwise, the pictogram was discarded. Average rankings among the selected pictograms were calculated based on the rankings given by the human judges; if average rankings were the same among two or more pictograms, variances were calculated to give higher ranking to the pictogram with lower variance. As a result, a ranked relevant pictogram set for 188 words were created and used in the evaluation.

[‡]There were initially a total of 903 unique pictogram interpretation words for 120 surveyed pictograms which could be used as word queries for the retrieval task. We first performed retrieval tasks with these 903 words using the four approaches to eliminate 399 words that returned the same result for all four approaches, since these words would be ineffective in discerning the four approaches' retrieval performance. Another 216 words which returned the same results for the three semantic relevance approaches were eliminated. 288 words remained as a result. Among the 288 words, words having more than nine candidate pictograms, similar words (e.g. *hen*, *rooster*), singular/plural words (e.g. *girl*, *girls*), and varied tenses (e.g. *win*, *winning*) were eliminated leaving 188 words to be judged for relevancy. The constitution of major pictogram categories in the 903 words and 188 words were:
-903:[Agent,10%],[Matter,24%],[Event,61%],[Location,2%],[Time,3%]
-188:[Agent, 9%],[Matter,28%],[Event,50%],[Location,9%],[Time,5%]

[§]Candidate pictograms contain given interpretation.

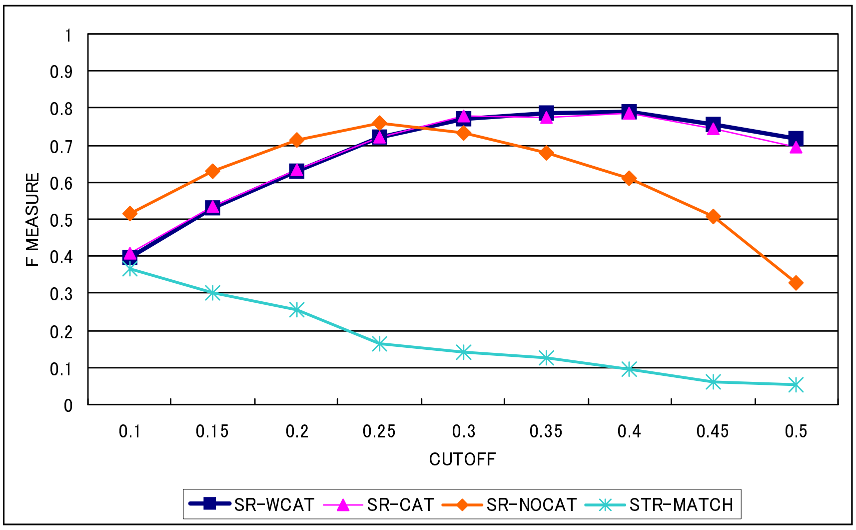


Figure 3.2: Comparison of F_1-measure graph of the four approaches

3.4.3 Precision, Recall, and F_1 measure

The mean precision, mean recall, and F_1 measure [Rijsbergen 79] of 188 retrieval tasks on the four pictogram retrieval approaches were calculated using nine different cutoff points from 0.1 to 0.5 with 0.05 intervals. Figure 3.2 shows F_1 measure, and Tables 3.3 and 3.4 respectively show the mean precision and mean recall: SR-WCAT indicates categorized and weighted semantic relevance approach; SR-CAT indicates categorized semantic relevance approach; SR-NOCAT indicates not-categorized semantic relevance approach; and STR-MATCH indicates the baseline string match approach. Note that the mean precision values were calculated using the valid tasks that returned at least one result. For example, in the case of cutoff value 0.5, only nine retrieval tasks returned at least one pictogram for the STR-MATCH approach; hence, the mean precision of the STR-MATCH approach was calculated using only those nine tasks. Note that gain in retrieval performance is achieved through semantic relevance and word categorization, but minimal gain is obtained through word weighting.

Table 3.3: Mean precision of four approaches at different cutoffs

CUTOFF	SR-WCAT	SR-CAT	SR-NOCAT	STR-MATCH
0.10	0.24850	0.25810	0.34883	0.98056
0.15	0.36089	0.36467	0.46397	0.99275
0.20	0.46108	0.46512	0.57565	1.00000
0.25	0.57463	0.57928	0.67917	1.00000
0.30	0.65529	0.66786	0.73870	1.00000
0.35	0.70685	0.70442	0.79100	1.00000
0.40	0.73910	0.74880	0.84497	1.00000
0.45	0.76704	0.76979	0.87760	1.00000
0.50	0.78753	0.81036	0.89655	1.00000

Table 3.4: Mean recall of four approaches at different cutoffs

CUTOFF	SR-WCAT	SR-CAT	SR-NOCAT	STR-MATCH
0.10	1.00000	1.00000	0.99867	0.22615
0.15	1.00000	0.99823	0.97442	0.17766
0.20	0.99493	0.98980	0.94174	0.14752
0.25	0.96226	0.95713	0.86184	0.08901
0.30	0.94125	0.93784	0.72376	0.07704
0.35	0.88712	0.86527	0.59734	0.06640
0.40	0.84705	0.82768	0.47810	0.05044
0.45	0.74785	0.72214	0.35887	0.03183
0.50	0.65888	0.60657	0.20222	0.02739

3.5 Discussions

We see in Figure 3.2 that a broader cutoff range between 0.24 and 0.5 is obtained by the categorized approach for F_1 measure greater than 0.7 (SR-WCAT & SR-CAT) whereas the not-categorized approach has a more steeper curve with narrower cutoff range between 0.19 and 0.33 (SR-NOCAT). The wider range of stable F_1 measures given by the categorized approach owes to a priori grouping of the interpretation words into related perspectives; this enables targeted semantic relevance calculation on words more related to the query and related to each other leading to the improvement in recall without damaging precision. This is confirmed in Tables 3.3 and 3.4: in Table 3.4, the recall range of SR-WCAT and SR-CAT is tighter with the range approximately between 0.6 and 1.0 whereas SR-NOCAT is broader with recall range approximately between 0.2 and 1.0; meanwhile in Table 3.3, the precision range of all three approaches, SR-WCAT, SR-CAT, and SR-NOCAT, are similar with SR-WCAT and SR-CAT approximately in between 0.25 and 0.8, and SR-NOCAT in between 0.35 and 0.9.

The fact that no significant performance gain was obtained through category weighting of the words should be discussed. The categorical semantic relevance values of SR-WCAT and SR-CAT did not differ greatly in most cases of the retrieved results (in general, SR-WCAT had slightly higher values than SR-CAT). As a result, the same set of pictograms was retrieved for both approaches except in those cases where the two categorical semantic relevance values branched at the cutoff point. Analyzing the exceptional cases where large value differences between the two approaches were observed revealed that the combination of large category weight on the query word together with small category weight on the surrounding interpretation words triggered a drastic change in the interpretation word ratio, causing the categorical semantic relevance value of SR-WCAT to increase drastically. Such an exceptional case was rarely observed, however, and in most cases the constitution of the interpretation word ratio in SR-WCAT and SR-CAT were very similar. This is why the two approaches exhibited very similar retrieval performances.

Our method can be applied to various image management applications such as clipart search systems or online photo-sharing systems as long as the images are labeled with descriptive tags, and that those tags have frequencies; but the benefits of categorization can be more fully enjoyed through a novel application which generates a pictogram-mixed sentence. We will call it a *pictogram sentence generator*. A pictogram sentence generator is a parser-like application which takes a text sentence as input and outputs a pictogram-mixed sentence. The generator first parses the text sentence to generate a parse tree, and then takes the lemma of the word in the tree to use it as a word query to search for the most relevant pictogram to replace the word. The pictogram retrieval system introduced in section 3.3.5 is utilized in the search process, but instead of ranking the retrieved result using the major pictogram category of the query, the generator specifies which pictogram category to emphasize, i.e. which categorical semantic relevance value should be selected and ranked. We explain these using two examples. Suppose we want to convert the following two sentences, both containing the word "slide", into a pictogram-mixed sentence:

(1) John likes to slide down the hill.

(2) John played on the slide.

It is obvious to humans that the word "slide" is used differently in the two sentences, but a simple string match by a machine will find no difference. One way to allow a machine to discern this kind of usage difference is to provide the machine with semantic role information. Recent advances in semantic role labeling technology have realized fairly accurate automatic labeling of semantic roles , and currently several semantic role labelers have been implemented. Some labelers uses semantic roles defined in the Proposition Bank [Palmer 05], and the numbered arguments in the frameset are aligned to VerbNet thematic roles. If we can map these thematic roles, adjuncts labels, verb information to the first, second, and third level classifications in the Concept Dictionary of the EDR Electronic Dictionary, then once the semantic roles are identified, we can obtain the first level classification in the EDR (i.e. the most appropriate pictogram category) that can be used for ranking the retrieved pictograms.

Going back to the two example sentences, the semantic role labeler will output "verb" as label for the word "slide" in sentence (1) and "AM-LOC location" for "slide" in (2); then, we can map "slide: verb" to `30f83e action/act` in the Concept Dictionary to obtain the EVENT category whereas map "slide: location" to `30f751 location/locale/place` to obtain the LOCATION category. When the retrieved pictograms are ranked using the acquired pictogram categories, the generator will output Figure 3.1's top-most pictogram showing a person sliding for the word "slide" in sentence (1) and the bottom-most pictogram showing a playground slide for the word “slide” in (2).

The number of the human judges participated in the evaluation experiment given in this paper were few and the age group was limited, and so a greater number of human judges encompassing a wide range of age groups should be incorporated, and more diverse queries should be used for performance evaluation to accommodate real-world usage. Moreover, reliability metric to mitigate the overrated calculation of single word categorical semantic relevance value should be defined to improve precision in the future.

3.6 Summary

Polysemous and shared pictogram interpretation can lead to ambiguity in pictogram interpretation, which can cause misunderstanding in communication using pictograms. To retrieve pictograms that can better convey the intended meaning, we proposed a method of selecting and ranking relevant pictograms which are more likely to be interpreted as intended. We proposed a categorical semantic relevance measure, which calculates how relevant a pictogram is to a given interpretation in terms of a pictogram category. The measure defines the probability and similarity measurement of categorized pictogram interpretations. Five pictogram categories used for categorizing pictogram interpretation words were defined using the Concept Dictionary of the EDR Electronic Dictionary. Three semantic relevance approaches, (i) not-categorized semantic relevance approach, (ii) categorized approach, and (iii) categorized and weighted approach, were evaluated using five human judges and 188 queries, and the categorized approaches showed more stable performance than the not-categorized approach.

Chapter 4

Human Cultural Difference Detection Criteria

This chapter identifies which part of cross-cultural pictogram interpretations do humans pay attention to when detecting cultural differences in pictogram interpretations. A two-part human subject study consisting of a questionnaire answering and face-to-face interview was conducted to clarify human cultural difference detection criteria.

4.1 Introduction

Pictograms have clear pictorial similarities with some object [Kolers 69], and one who can recognize the object depicted in the pictogram can interpret the meaning associated with the object. In this chapter, we again take up the pictograms used in computer-mediated intercultural communication, specifically, those used in a children's email system [Takasaki 07]. The advantage of using pictograms in intercultural communication is that both human participants can attempt to communicate without using any second language. This is not the case when certain natural language is used in intercultural communication; at least one participant must communicate using a second language. Pictorial symbols, however, are not universally interpretable. For example, the cow is a source of nourishment to westerners who drink milk and eat its meat, but it is an object of veneration to many people in India. Hence, a picture of cow could be interpreted quite differently by Protestants and Hindus [Kolers 69].

When pictograms having culturally different interpretations are used in intercultural communication, misunderstanding may arise between participants having different cultural backgrounds. There can be several solutions to this interpretation problem: one is to design a set of pictograms that have unique and clear meaning; the other is to switch the pictogram selected by the sender to a more culturally appropriate pictogram suited to the receiver. To do this, we need to detect pictograms having culturally different interpretations.

Detecting cultural differences requires an understanding of culture, and one way to understand culture is to study existing definitions on culture. More than hundred definitions on culture exists [Kroeber 52], but here we look at Geertz's definition: he defines culture as "a historically transmitted pattern of meanings embodied in symbols, a system of inherited conceptions expressed in symbolic form by means of which men communicate, perpetuate, and develop their knowledge about and attitudes towards life [Geertz 73]". Based on this definition, cultural differences can be viewed as differences in the 'pattern of meanings' or semantic differences; hence, detecting cultural differences can be viewed as detecting semantic differences.

Existing computational methods for calculating semantic differences (or dissimilarity) in two documents make use of, for example, the vector space model [Manning 99] and hierarchical semantic nets [Rada 89]. In the vector space model, documents are represented in a high-dimensional space, in which each dimension of the space corresponds to a word in the document collection. Spatial proximity or closeness of the two documents is calculated by looking at the angles of the two documents' vectors [Manning 99]. On the other hand, in the hierarchical semantic network, the conceptual distance between any two concepts is defined as the shortest path through a semantic network of hierarchical relations [Rada 89]. Mapping the words in documents to concepts in the semantic network will enable word-distance calculation.

Applying such semantic dissimilarity calculation to the detection of cultural differences in pictogram interpretations could easily be envisaged, but whether the existing computational methods are sufficient in detecting cultural differences needs to be carefully studied. The goal of this paper is to clarify whether human detection of cultural differences in pictogram interpretations can be approximated using the existing methods of semantic dissimilarity calculation. To do this, we need to understand how humans detect cultural differences in pictogram interpretations.

This chapter reports the findings on a human cultural difference detection study conducted using cross-cultural pictogram interpretations as stimuli. Six human subjects participated in a two-part study consisting of (1) answering a questionnaire and (2) responding to a post-questionnaire interview. Based on the findings on how humans detect cultural differences, we will show that existing approaches of semantic dissimilarity calculation are insufficient in detecting a complete range of cultural differences in pictogram interpretations.

4.2 U.S.–Japan Pictogram Web Survey

4.2.1 Construction of Stimuli

A pictogram web survey was conducted in the U.S. and Japan to collect cross-cultural pictogram interpretations with possible cultural differences. The two countries were selected for their cultural distinctness and ease of data gathering and analysis.

Pictograms

120 pictograms used in a children's email system [Takasaki 07] were used as visual stimuli for the pictogram web survey. These pictograms were designed by Japanese art major college students who were novices at pictogram design. A list of words, which include Ogden's Basic English

words*, was given to these students, and they were asked to draw pictograms depicting each word. Unlike road signs designed by professionals for unambiguous interpretation, the design of these pictograms allowed culturally-dependent [Cho 07a], polysemous [Cho 08a] interpretations.

Survey Method

A free-answer pictogram web survey† was conducted in the U.S. and Japan via the World Wide Web from October 1, 2005 to November 30, 2006 (a total of 14 months). Using the 120 pictograms as visual stimuli, human respondents in the U.S. and Japan were asked to type how each pictogram can be interpreted inside the text field provided below each pictogram. A set of 10 pictograms was shown per webpage, and the respondents could answer as many question sets they liked. The maximum number of question sets a respondent could answer was 12 sets which contained a total of 120 pictograms.

Survey Data

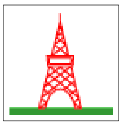
A total of 543 respondents in Japan and 935 respondents in the U.S. participated in the pictogram web survey. An average of 147 and 97 interpretations consisting of English and Japanese words or phrases respectively were collected for each pictogram (duplicate expressions included). For each country, unique interpretation words or phrases were listed for each pictogram, and the frequencies of those unique words were counted.

An example of U.S.–Japan word count result for one of the pictogram surveyed is shown in Table 4.1. The two columns on the left show interpretation words and frequencies collected from the U.S. respondents. The two columns on the right show those of the Japanese respondents. U.S. inter-

*66 words were selected from Ogden's Basic English words which include words on operations (e.g., here, there, tomorrow, yesterday), general things (e.g., man, money, place, sleep), picturable things (e.g., baby, bridge, nose, prison), and qualities (e.g., angry, fat, long, wrong).

†`http://www.pangaean.org/iconsurvey/`

Table 4.1: Example of tallied U.S.–Japan pictogram interpretations



U.S. Interpretations	Freq.	%	Japan Interpretations	Freq.	%
Eiffel Tower	25	19.2	Tokyo Tower (東京タワー)	41	44.6
Paris	25	19.2	tower (タワー)	22	23.9
tower	20	15.4	tower (塔)	7	7.6
Eifel Tower	6	4.6	Eiffel Tower (エッフェル塔)	6	6.5
France	6	4.6	tower (電波塔)	3	3.3
radio tower	3	2.3	TV tower (テレビ塔)	2	2.1
travel	3	2.3	Tokyo (東京)	2	2.1
building	2	1.5	Paris (パリ)	1	1.1
oil rig	2	1.5	steel tower (鉄塔)	1	1.1
cell phone tower	1	0.8	tower (電波搭)	1	1.1
electricity tower	1	0.8	Eiffel Tower (えっふぇるとう)	1	1.1
fun/expensive	1	0.8	tower (たわー)	1	1.1
oil	1	0.8	Tokyo Tower (とうきょうたわー)	1	1.1
technology	1	0.8	Tokyo Tower (とうきょうたわ〜)	1	1.1
tour	1	0.8	tower (ターわー)	1	1.1
omitted words	32	24.6	tower & Tokyo (タワー (東京でもあり))	1	1.1
Total	130	100	Total	92	100

Note: Omitted words for the U.S. interpretations include: 24 misspelled "Eiffel Tower," four geographic names such as "Paris/Tokyo", one "birge", one "monument", one "Pisa", and one "tourist sights".

pretation word "Eiffel Tower" placed at the top has a frequency of 25, and this means that twenty-five U.S. respondents wrote "Eiffel Tower" as the interpretation for the pictogram. Several low frequency U.S. interpretation words were omitted to save space.

U.S.–Japan pictogram interpretation tables similar to that of Table 4.1 were created for all 120 pictograms. These tables contained the original Japanese interpretation words, but in this paper we translate all Japanese words into English for readability. Two Japanese-English dictionaries, 'EDICT[‡]' and 'Eijiro on the Web[§]', were used as references.

[‡] `http://www.csse.monash.edu.au/~jwb/j_edict.html`
[§] `http://www.alc.co.jp/`

4.2.2 Characteristics of U.S.–Japan Pictogram Interpretations

Analyzing the 120 U.S.–Japan pictogram interpretations revealed the following characteristics in the pictogram interpretations. We explain each characteristic by giving examples using the interpretation words in Table 4.1:

- Geographically and/or psychologically proximate real-world subject matter is identified (e.g., “Tokyo Tower”, “Tokyo”, “Eiffel Tower”, “Paris”)

- The position of interpretation word in the concept hierarchy is varied (e.g., “radio tower”, “cell phone tower”, “electricity tower” are subconcepts of “tower” while “building” is a superconcept of “tower”. By contrast, “Eiffel Tower” is an instance of “tower”.)

- Concrete object(s) depicted in the pictogram is associated with some abstract concept (e.g., “Tower” is associated with “travel”, “tour”, or “technology”.)

- Most pictograms have polysemous interpretations (e.g., “Eiffel Tower”, “tower”, “Paris”, “travel”, “oil”)

When each country’s pictogram interpretation words are considered as a set of concepts, the relationship between U.S.-Japan pictogram interpretation words can be classified into one of the four relations given below:

Disjoint Relation Almost no common concept exists between U.S. and Japan.

Intersecting Relation U.S. and Japan share some common concepts, but dissimilar concepts exist in both countries.

Subset Relation U.S. concepts subsume Japanese concepts or vice versa.

Equivalent Relation Most U.S. and Japanese concepts are the same.

It is natural to think that no cultural difference exists if most of the two countries' interpretations are similar, i.e., the two countries' concepts have an equivalent relation. The classification of relation, however, is dependent on the level of interpretations being handled: the interpretations can be handled either at the word-level or at the concept-level. For instance, U.S.–Japan interpretations in Table 4.1 can be classified as having an intersecting relation when viewed at the word-level since they share the word "tower", but also have dissimilar interpretations such as "Eiffel Tower" and "Tokyo Tower". By contrast, when viewed at the concept-level, both countries' interpretations are centered on the concept of "tower", so they can be classified as having an equivalent relation. In the former case, the two countries' interpretations may be viewed as having cultural differences whereas in the latter case they can be viewed as having no cultural difference.

4.2.3 Selection of Stimuli for Human Detection Study

With both possibilities (presence/absence of cultural differences) mixed in the two countries' pictogram interpretations, how do humans detect cultural differences? A two-part study was designed to clarify this: the first part of the study asks human subjects to assess cultural differences using a questionnaire; the second part of the study interviews human subjects for the reasons behind their assessment (i.e., cultural difference detection criteria). To conduct this study, U.S.–Japan pictogram interpretations with greater cultural differences were selected as stimuli. They were selected as follows:

Three human judges independently analyzed the 120 U.S.–Japan pictogram interpretation tables for cultural differences; the main criterion for determining cultural differences was the presence of semantic differences in the two countries' interpretations. A total of 60 pictograms were judged to have cultural differences by the three judges: of these, 19 pictograms were found to have culturally different interpretations by two or more judges, and 7 pictograms were found to have culturally different interpretations by all three judges. We use 60 pictograms with possible cultural differences as stimuli for the human cultural detection study described next.

4.3 Human Detection of Cultural Differences

Shorter lists of interpretation words were created from the 60 U.S.–Japan pictogram interpretations by merging misspelled words, fluctuation in expressions, and synonymous words. Moreover, to reduce the time and effort of the human subjects participating in the main study, a preliminary study was conducted to further select 30 pictogram interpretations with greater cultural difference. These pictogram interpretations were used in the main study which consists of a questionnaire and interview.

4.3.1 Preliminary Study

Four human subjects (different from the three judges who previously selected 60 pictogram interpretations out of 120) were asked to assess cultural differences of the 60 U.S.–Japan pictogram interpretations by answering a questionnaire. The questionnaire used in the preliminary study was similar to the one shown in Figure 4.1. Each subject was instructed to choose a level of cultural difference using a seven-point Likert scale: the levels given were "Strong cultural difference exists", "Cultural difference exists", "Cultural difference exists somewhat", "Undecided", "Rather no cultural difference", "No cultural difference", and "Absolutely no cultural difference". After the four subjects answered the questionnaire, the levels were converted to numerical values ranging from 7 (Strong cultural difference exists) to 1 (Absolutely no cultural difference). The four human subjects' cultural difference assessments of the 60 pictogram interpretations were averaged and the top 30 pictogram interpretations with greater cultural differences were selected as stimuli for the main study.

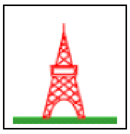
4.3.2 Main Study

Subjects & Method

Six human subjects (different from the three judges and four subjects in the previous studies) participated in the cultural difference detection study.

28. Below are the U.S.-Japan interpretations for the pictogram placed in the middle.

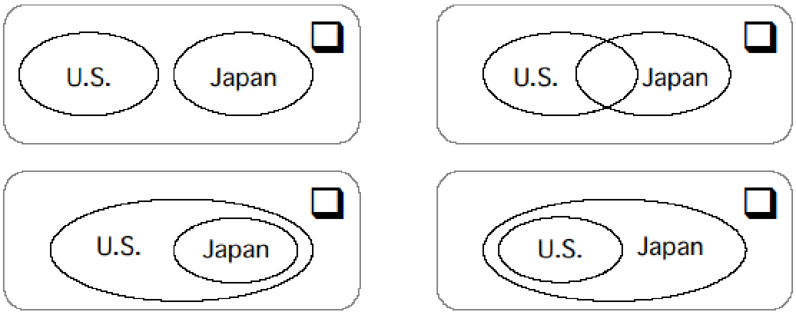
U.S. Interpretation	
Eiffel Tower	48%
Paris	24%
tower	17%
France	7%
radio tower	2%
travel	2%



Japanese Interpretation		
東京タワー	47%	Tokyo Tower
タワー，塔	35%	tower, tower
エッフェル塔	8%	Eiffel Tower
電波塔	5%	radio tower
東京	3%	Tokyo
テレビ塔	2%	TV tower

28-1. Circle all the similar interpretations (common interpretations).

28-2. Which diagram most closely represents the relationship between the two countries' pictogram interpretations? Check one box.



28-3. Based on the two interpretations and their ratios, similar interpretations, and the two interpretations' relationship, do cultural differences exist? Check one box.

- Strong cultural difference exists ☐
- Cultural difference exists ☐
- Cultural difference exists somewhat ☐
- Undecided ☐
- Rather no cultural difference ☐
- No cultural difference ☐
- Absolutely no cultural difference ☐

Figure 4.1: Questionnaire given to the U.S. subjects in main study

The two-part study consisted of (1) answering a questionnaire (Figure 4.1) and (2) responding to a post-questionnaire interview. Three subjects were U.S. nationality English teachers living in Japan for over five years with a fair understanding of Japanese.

Other three subjects were Japanese graduate students with graduate-school level English knowledge. All six subjects were paid for their participation. Figure 4.1 shows a sample page of the English questionnaire given to the U.S. subjects. (Japanese subjects were given a Japanese questionnaire.) During the questionnaire-answering part of the study, the subjects were instructed to first mark the two countries' interpretations for similar interpretations, and then select one of the four Venn diagrams which depicted the relationship between the two countries' interpretations: (i) disjoint relation, (ii) intersecting relation, (iii) subset relation in which U.S. interpretation subsumes Japanese interpretation, and (iv) subset relation in which Japanese' subsumes U.S.'s.

Based on the two countries' pictogram interpretations and their percentages, similar interpretations, and the relationship between the two countries' interpretations, the subjects assessed the level of cultural difference in each U.S.–Japan pictogram interpretations. The same seven-point Likert scale used in the preliminary study was used in the main study. After the questionnaire was answered, a one-hour interview asking the reasons behind the cultural difference assessment was conducted to each subject. We first present the results of the cultural difference assessment questionnaire.

Questionnaire Result

Table 4.2 shows cultural difference assessment values of the six human subjects: columns A1, A2, A3 in Table 4.2 show cultural difference assessment values given by three U.S. subjects; columns J1, J2, J3 show values given by three Japanese subjects. The same numerical conversion used in the preliminary study (7 for "Strong cultural difference exists" and 1 for "Absolutely no cultural difference") were used. The values are sorted with greater average cultural difference assessment value on top.

Table 4.2: Six humans' cultural difference assessment results

PIC.	A1	A2	A3	J1	J2	J3	AVG	SD
P21	7	7	7	7	7	7	7.00	0.00
P01	7	7	7	7	7	6	6.83	0.37
P12	6	7	7	6	5	7	6.33	0.75
P28	6	6	6	6	6	7	6.17	0.37
P02	6	6	7	6	7	5	6.17	0.69
P14	6	6	6	6	5	7	6.00	0.58
P11	6	7	3	6	7	7	6.00	1.41
P13	6	7	7	5	3	7	5.83	1.46
P15	5	7	7	5	4	6	5.67	1.11
P10	5	7	7	6	2	6	5.50	1.71
P16	5	6	5	6	3	6	5.17	1.07
P30	5	5	5	5	4	6	5.00	0.58
P22	5	5	3	5	5	5	4.67	0.75
P08	5	6	6	5	3	3	4.67	1.25
P09	3	5	6	5	3	6	4.67	1.25
P23	5	3	5	5	5	4	4.50	0.76
P07	5	5	7	2	4	2	4.17	1.77
P03	2	5	5	5	2	5	4.00	1.41
P04	2	5	5	5	2	5	4.00	1.41
P05	2	5	2	6	3	6	4.00	1.73
P18	2	5	6	5	2	3	3.83	1.57
P20	2	3	5	2	6	4	3.67	1.49
P17	2	7	2	6	3	2	3.67	2.05
P06	3	5	3	3	2	5	3.50	1.12
P19	3	5	5	3	2	3	3.50	1.12
P27	2	6	2	5	2	4	3.50	1.61
P29	4	3	1	5	2	5	3.33	1.49
P25	3	3	3	3	2	5	3.17	0.90
P26	3	1	3	2	3	5	2.83	1.21
P24	2	3	3	2	2	2	2.33	0.47

Note: A1-A3 indicate U.S. subjects; J1-J3 indicate Japanese subjects; the numbers indicate: 7 = Strong cultural difference exists, 6 = Cultural difference exists, 5 = Cultural difference exists somewhat, 4 = Undecided, 3 = Rather no cultural difference, 2 = No cultural difference, and 1 = Absolutely no cultural difference.

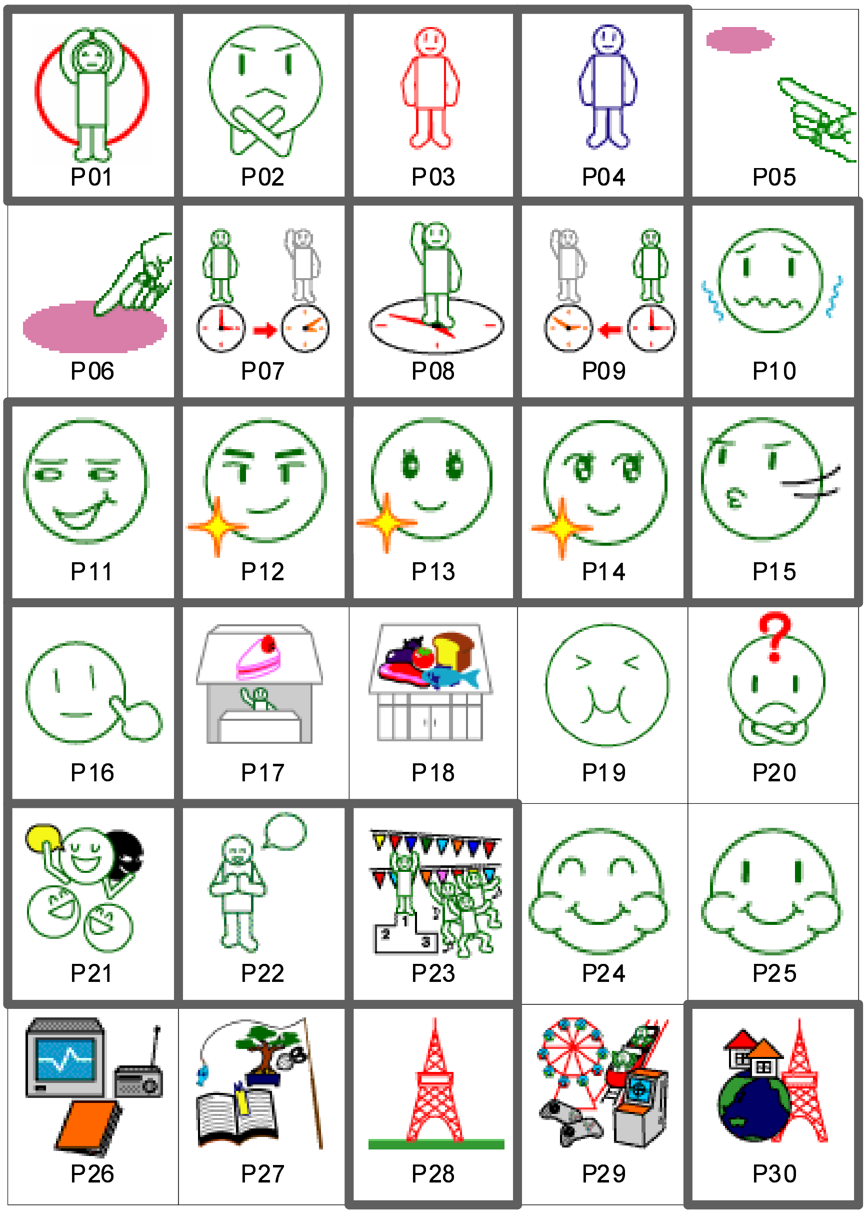


Figure 4.2: Thirty pictograms questioned for cultural differences (Bold-lined pictograms are assessed to have some cultural differences)

The images of the 30 pictograms are shown in Figure 4.2. Each row in Table 4.2 shows six human subjects' cultural difference values (A1-J3), six subjects' average assessment value (AVG), and the standard deviation (SD) of each pictogram. We interpreted four or more human subjects assessing average assessment values of 4 or greater (AVG $\geq$ 4) to mean some kind of cultural difference exist; this is because 4 is the mid-value in the cultural difference assessment scale. A total of 6 pictograms (P21, P01, P12, P28, P02, P14) were unanimously assessed by six subjects to have some kind of cultural difference (all six subjects returned values 5 or greater). A total of 19 pictograms, including the unanimously assessed 6 pictograms mentioned just now, were assessed to have some kind of cultural difference (AVG $\geq$ 4, four or more people returned assessment of 4 or greater). The 17 pictograms were all included in the 19 pictograms that were initially assessed to have cultural differences by two or more human judges. The two pictograms that were not included were P05 and 06. P21 had the highest average cultural difference value of 7.00.

Post-Questionnaire Interview

Face-to-face interviews were conducted to the six subjects to elucidate the reasons behind their cultural difference assessment. The following aspects in U.S.–Japan pictogram interpretations were taken into account when assessing cultural differences:

- Similar/dissimilar interpretations in the two countries
- Percentage or ranking of the interpretations
- Conformity/variance of semantics within one country's interpretations
- Presence of proper nouns (e.g. country names)
- Positive/negative connotation in the interpretations

Table 4.3: Thirty pictogram data: P01–P04 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P01 (AVG: 6.83, RANK: 2)			
exercise, exercising	46	okay (OK, オッケー)	44
stretch, stretching	14	circle (まる, 丸, 〇)	21
jump rope	11	correct answer (正解)	14
happy	7	all rightie (いいよ)	11
circle, circular	7	yes (はい)	5
yoga	5	bingo (あってる, あたり)	3
dance	5	good (良い)	2
jumping jacks	5		
P02 (AVG: 6.17, RANK: 5)			
mad	44	no-no (だめ)	59
angry	39	penalty (ばつ)	19
no	6	no (いいえ)	5
stubborn	4	not a chance (やだ)	4
anger	3	wrong answer (不正解)	4
frustrated	2	miss (はずれ)	4
upset	2	think (考える)	3
		No Good (NG)	2
P03 (AVG: 4.00, RANK: 18)			
woman, female	37	lady, female, woman (女の人, 女性, 女)	78
man	21	mom (お母さん)	9
mom, mother	17	adult female (大人の女性)	8
dad, father	11	adult (大人)	3
adult	9	person (人)	2
person	3		
adult woman	1		
teenager	1		
P04 (AVG: 4.00, RANK: 19)			
man, male	46	man, male (男の人, 男性, 男)	74
dad, father	17	adult man (大人の男性)	12
woman	12	dad (お父さん)	10
mom, mother	9	adult (大人)	2
adult	9	person (人)	2
person	3		
boy	2		
adult man	2		

Table 4.4: Thirty pictogram data: P05–P08 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P05 (AVG: 4.00, RANK: 20)			
up, above	30	there (あそこ)	36
there, over there, go there	30	that (あれ)	31
point, pointing	13	over there (あっち, あちら)	14
far, away	12	above (上)	8
high	5	far (遠い, 遠く)	8
hole	4	cloud (雲)	3
circle	3		
small	3		
P06 (AVG: 3.50, RANK: 24)			
down, pointing down	34	here (ここ)	41
here, right here	28	this (これ)	30
near, close	12	close, near (近い, 近く)	8
touch, touching	7	down (下)	7
big hole	5	this direction (こっち)	6
low	5	puddle (水たまり)	4
circle	5	manhole (マンホール)	2
big	4	ground (地面)	2
P07 (AVG: 4.17, RANK: 17)			
late	31	future (未来)	36
time	24	time passes (時間が経つ)	17
10 minutes	10	after 10 minutes (10 分後)	16
later	8	afterward (あとで)	9
leave	7	be late (遅刻)	9
future	6	time (時間)	5
on time	6	bye-bye (バイバイ)	5
after	4	appointment (待ち合わせ)	3
boy and girl	4		
P08 (AVG: 4.67, RANK: 14)			
on time	31	now, present (今, 現在)	46
time	22	time (時間)	19
now, present	10	the time now (今の時間)	9
hi, hello	9	appointment (待ち合わせ)	7
time to go	9	compass (方位磁石)	5
what time is it	7	watch (時計)	5
clock	6	just (ちょうど)	5
waiting	6	on time (時間通り)	4

Table 4.5: Thirty pictogram data: P09–P12 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P09 (AVG: 4.67, RANK 15)			
early	34	past (過去)	42
10 minutes ago	13	10 minutes before (10 分前)	21
before	11	time is turned back (時間が戻る)	20
past	10	before (前)	12
time	10	a little while ago (さっき)	5
late	8		
10 minutes	7		
earlier	7		
P10 (AVG: 5.50, RANK: 10)			
scared, afraid	29	cold (寒い)	64
worried, worry	17	scared (怖い)	27
cold	16	shiver (震える)	7
nervous	15	feel chilly (ぞくぞく)	2
sad	13		
confused	7		
sick	3		
P11 (AVG: 6.00, RANK: 7)			
sly, cunning, slick	26	sneer, grin, bitter smile (嘲笑, にやり, 苦笑い)	22
sneaky	25	hehehe (laughing sound) (へへへ)	20
smile, smirk, snicker	13	be doubtful, doubt (疑う, 疑い)	20
happy	12	make a fool of (ばかにする)	18
shy, coy	7	interesting (おもしろいな)	7
flirting	5	make fun of (からかう)	5
cool	4	scheme (たくらむ)	4
snide	4	slander (悪口)	4
funny, silly	4		
P12 (AVG: 6.33, RANK: 3)			
mischievous	20	cool (かっこいい)	59
happy	16	handsome (ハンサム)	22
smart	14	boast (自慢)	12
mean	13	dazzle (キラーン)	7
handsome	11		
boy	11		
cool	8		
evil	7		

Table 4.6: Thirty pictogram data: P13–P16 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P13 (AVG: 5.83, RANK: 8)			
happy	43	cute, pretty (かわいい, きれい)	76
pretty, cute, beautiful	21	beautiful woman, beauty (美人, 美女)	11
nice, good	10	ufufu (laughing sound) (うふふ)	5
girl	9	smiling face (笑顔)	4
sweet	8	dazzle (キラキラ)	4
innocent	6		
proud	3		
P14 (AVG: 6.00, RANK: 6)			
happy	24	pretty, cute (きれい, かわいい)	47
cute, pretty, beautiful	21	beautiful woman (美人)	24
in love	13	dazzle (キラキラ)	7
sweet	13	ufufu (laughing sound) (うふふ)	7
flirting	11	great (すてき)	6
shy, coy	7	prim (おすまし)	6
thinking	6	feminine (女らしい)	3
thoughtful	5		
P15 (AVG: 5.67, RANK: 9)			
whistling, whistle	43	nonchalant, ignore (知らんぷり, 無視する)	38
no	13	humph (ふん)	19
blow, blowing	8	pout (すねる)	14
ignore, ignoring	7	pui (abruptly turning) (ぷい)	10
kiss, kissing	7	dislike (嫌い)	8
surprised	7	whistle (口笛)	8
annoyed	7	displeasure (不機嫌)	3
indifference	4		
wind, windy	4		
P16 (AVG: 5.17, RANK 11)			
thinking, idea	24	I, me, myself (わたし, ぼく, 自分)	61
face	16	cheek (頬)	13
me	14	face (顔)	9
cheek	13	mirror (鏡)	7
point, pointing	12	you (あなた)	5
touch	6	photograph (写真)	3
eye	5	smile (笑顔)	2
happy	5		
you	5		

Table 4.7: Thirty pictogram data: P17–P20 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P17 (AVG: 3.67, RANK: 23)			
cake	29	cake shop (ケーキ屋)	93
pie	22	cake (ケーキ)	7
bakery	18		
dessert	15		
cake shop	7		
dessert store	5		
pastry store	4		
P18 (AVG: 3.83, RANK: 21)			
food	43	supermarket (スーパー)	87
grocery store	31	grocery store (食料品店)	7
food groups	8	food (食べ物)	4
market	7	greengrocer (八百屋)	2
meal	5		
supermarket	3		
restaurant	3		
P19 (AVG: 3.50, RANK: 25)			
sour	53	sour (酸っぱい)	48
full	14	delicious (おいしい)	36
happy	12	full stomach (満腹)	6
eating	9	funny (おかしい)	4
chew, chewing	8	happy (嬉しい)	2
mouthful	4	fun (たのしい)	2
		suppress one’s laughter (笑いをこらえる)	2
P20 (AVG: 3.67, RANK: 22)			
confused, puzzled	73	think (考える)	48
thinking	8	question (疑問)	17
question	7	Why is that? (なんでだろう)	13
I don’t know	6	What is it? (なんだろう)	10
wondering	3	Let me see··· (はてな)	6
I don’t understand	3	wonder (不思議)	6

Table 4.8: Thirty pictogram data: P21–P24 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P21 (AVG: 7.00, RANK: 1)			
talking, conversation	20	to lie, liar, lie (嘘をつく, 嘘つき, 嘘)	42
friends	19	double-dealing (裏表)	20
party	17	deceive (騙す)	12
joking	11	scheme (何かをたくらむ)	8
gossip	9	dual personality (二重人格)	6
laughing	9	boast, big talk (自慢, 自慢話)	6
happy group	8	backbiting (陰口)	6
happy	7		
P22 (AVG: 4.67, RANK: 13)			
talking, speaking	33	speak (話す)	23
praying, pray	21	announcement (発表)	17
thinking	16	thank you (ありがとう)	15
talking to oneself	8	soliloquy (独り言)	13
lonely	7	please (お願い)	11
reading	6	pray (祈る)	7
singing	5	I tell you what. (そうだ)	6
introduce oneself	4	feel relieved (ほっとする)	4
		scenario reading (本読み)	4
P23 (AVG: 4.50, RANK: 16)			
winner, champion	49	athletic event (運動会)	45
win, winning, victory	23	No. 1, first place (一番, 一位)	28
event	10	victory (優勝)	13
first place	7	first place prize (一等賞)	8
cheering	4	be glad (よろこぶ)	6
celebration	4		
award	3		
P24 (AVG: 2.33, RANK: 30)			
chewing, chew	42	delicious (おいしい)	67
happy	14	eat (食べる)	10
yummy	13	fun, happy (楽しい, 嬉しい)	9
full	12	chew, mumble (かむ, もぐもぐ)	8
mouthful	10	contentment (満足)	3
eating	5	smile (笑う)	3
sleeping	4		

Table 4.9: Thirty pictogram data: P25–P28 (continued next page)

U.S. Interpretations	%	Japan Interpretations	%
P25 (AVG: 3.17, RANK: 28)			
happy	31	delicious (おいしい)	24
chewing	18	eat (食べる)	21
yummy	17	so-so (まあまあ)	16
mouthful	8	mumble (もぐもぐ)	12
full	7	stuff one's mouth (ほおばる)	10
eat	6	full (いっぱい)	7
good	5	ordinary taste (ふつうの味)	6
like	4	happy (うれしい)	4
good food	4		
P26 (AVG: 2.83, RANK: 29)			
TV	18	information (情報)	35
media	15	media (メディア)	22
radio	15	radio (ラジオ)	15
entertainment	13	electric wave (電波)	6
electronics	13	communication (通信)	5
technology	8	electrical appliances (電化製品)	5
information	8	means of communication (通信手段)	4
communication	5	television (テレビ)	4
computer	5	news (ニュース)	4
P27 (AVG: 3.50, RANK: 26)			
hobbies	19	hobby (趣味)	87
fishing	16	play (遊び)	5
activities	15	fishing (つり)	3
reading	11	outdoor (アウトドア)	1
relax	11	book (本)	1
leisure	10	comic (マンガ)	1
outdoors	10	garden (庭)	1
books	8	leisure (レジャー)	1
P28 (AVG: 6.17, RANK: 4)			
Eiffel Tower	48	Tokyo Tower (東京タワー)	47
Paris	24	tower (タワー, 塔)	35
tower	17	Eiffel Tower (エッフェル塔)	8
France	7	radio tower (電波塔)	5
radio tower	2	Tokyo (東京)	3
travel	2	TV tower (テレビ塔)	2

Table 4.10: Thirty pictogram data: P29–P30

U.S. Interpretations	%	Japan Interpretations	%
P29 (AVG: 3.33, RANK: 27)			
carnival	31	amusement park (遊園地)	69
amusement park	20	amusement, play (遊び, 遊ぶ)	19
games	15	entertainment (娯楽)	6
fun	13	leisure (レジャー)	4
fair	13	Tokyo Disneyland (東京ディズニーランド)	2
entertainment	4		
amusement	4		
P30 (AVG: 5.00, RANK: 12)			
world	45	world (世界)	26
Paris, France	24	earth (地球)	24
travel	17	electric wave (電波)	10
earth, globe	8	Internet, network (インターネット, ネットワーク)	10
community	6	communication (コミュニケーション)	10
		country, Japan, Tokyo (国, 日本, 東京)	10
		Eiffel Tower (エッフェル塔)	6
		house (家)	4

With regard to the top nineteen pictograms with cultural differences (Table 4.2), they were assessed to have cultural differences for the following reasons:

1. Few similar interpretations exist between the two countries. (All pictograms except P28)

2. Quite a few similar interpretations exist, but the percentages of those interpretations are different between the two countries. (P01, P02, P04, P07, P10, P11, P12, P13, P14, P15, P16, P22, P23)

3. Conformity of semantics is observed in one country's interpretations while variance is observed in the other. (P10, P12, 14)

4. Proper nouns such as the name of a country or city exist in the interpretations. (P28)

5. Negative connotation in the interpretations is observed in one country while positive connotation is observed in the other. (P11, P12, P15, P21)

Except for P28, more than one reason was mentioned as cultural difference detection criteria in each of the top nineteen pictogram interpretations.

4.4 Discussions

The general criterion the human subjects used for assessing cultural differences was semantic similarity or dissimilarity between the two countries' interpretations. What human subjects considered as semantically similar, however, differed among subjects. For example, in the case of Figure 4.2 pictogram P29, the major U.S.–Japan pictogram interpretations were "carnival" (31%) and "amusement park" (69%) respectively (Table 4.2, P29), but they were perceived as different by two subjects (Table 4.2, P29: J1, J3) while three subjects perceived them as similar (Table 4.2, P29: A2, A3, J2). The former reasoned that "carnival" is an event while "amusement park" is a place, so the two were different; but the latter used association to conclude that the two countries' interpretations were similar.

If we try to approximate human perception of semantic similarity using the existing computational approaches, we find that it does succeed in some cases: for example, if we map P29's "carnival" and "amusement park" to SUMO ontology [Niles 01] (which is a kind of hierarchical semantic network) to calculate the conceptual distance, we find that "carnival" and "amusement park" are respectively mapped to RecreationOrExercise Class and Corporation Class which are very different classes; the conceptual distance between the two classes becomes very far, so the two interpretations are measured as very dissimilar leading to the conclusion that cultural difference exist in P29. But if we look at the average cultural difference assessment value of the six subjects, we find that P29 is assessed to have rather no cultural difference (Table 4.2, P29 AVG=3.33); hence, the hierarchical semantic net approach does not succeed in this case.

The semantic net approach also fails in the case of P28 for the opposite reason: P28's major U.S.–Japan interpretations are "Eiffel Tower" and "Tokyo Tower" respectively (Table 4.9, P28), and P28 is assessed to have cultural difference (Table 4.2, P28 AVG=6.17); but when the two interpretations are mapped to the SUMO ontology, they are mapped to the same Tower Class rendering their differences to disappear.

Such different perception on semantic similarity affects the degree of similar interpretations shared by two countries, and that in turn affects the relationship of the two countries' interpretations: for example, in the case of P29, if "carnival" and "amusement park" are considered as similar, the two countries' common interpretations will increase and the two countries' interpretation relationship will move toward equivalent relation; but when the two interpretations are considered as dissimilar, common interpretations will decrease, and the two countries' relationship will move toward disjoint relation.

Even if the two countries share similar interpretations, it may not guarantee the absence of cultural difference: for example, in the case of Figure 4.2 pictogram P10, "cold" and "scared" are the two major interpretations shared by the two countries (Table 4.3.2, P10), but the percentages and rankings of these interpretations are different, and the conformity vs. variance of semantics is observed in Japanese vs. U.S. interpretations respectively, leading to the assessment that cultural difference exists somewhat (Table 4.2, P10 AVG=5.50).

Using the vector space model to detect cultural differences in P10 will not succeed since structurally similar P27 (Table 4.9, P27) is assessed to have rather no cultural difference (Table 4.2, P27 AVG=3.50). The vector space model looks at the angle of two target vectors, and in this case the angle of two countries' interpretation vectors will be measured for cultural differences; since the structural constitutions of P10 and P27's U.S.–Japan interpretations are similar, they will probably yield somewhat similar interpretation vector pairs, resulting in similar angle values; but the two have opposite cultural difference assessments, and so using the vector space model to correctly detect cultural differences for both P10 and P27 will be difficult.

Other criteria for detecting cultural differences include the presence of contrasting pair of proper nouns (e.g., “Paris” vs. “Tokyo”) and negative/positive connotations underlying the two countries’ interpretations. As far as the finding in this study is concerned, these two factors seem to be independent to other factors which influence human cultural difference detection. There were yet other criteria for detecting cultural differences although mentioned infrequently. One subject pointed out that extreme conformity in interpretation was in itself a sign of cultural difference: Japanese interpretations in Table 4.9, P27 display this extremity. Another subject pointed out the differences in the focus of attention given to the object depicted in P27 in Figure 4.2 was a sign of cultural difference: in P27, Japanese focused on the overall object construction to arrive with “hobby” while U.S. respondents selectively focused on individual objects to come up with multiple concrete interpretations. The reasons for judging cultural differences in P27 differed for the two subjects (extreme conformity vs. focus of attention).

This hints us that simple yes or no cultural difference detection may be insufficient, and that the reasons behind the detection should be provided. Last but not least, human detection of cultural difference is also influenced by outside knowledge introduced by humans during the assessment: for instance, several subjects pointed out that Japan-specific gestures were interpreted in P1 and P2; idiosyncratic gender-color correlation was pointed out for P3 and P4; and Japanese traditional school athletic event was mentioned as the main reason for cultural difference in P23. Assuming that the two countries’ pictogram interpretations are provided in words, assigning humans to detect cultural differences in two countries’ pictogram interpretations is possible but not easy since it requires the human to have linguistic and cultural knowledge of both countries. Hence, machine detection of cultural differences may be helpful. With regard to the gathering of cross-cultural data, wide spread usage of tag-based applications nowadays enables us to gather human interpretations in the form of tags, and if user profile and/or IP address information can be utilized to categorize tags, we can realistically obtain cross-cultural interpretation data.

4.5 Summary

To understand whether human detection of cultural differences can be approximated using the existing computational approaches (for example, semantic similarity calculation), a study on how humans detect cultural differences in pictogram interpretations was conducted. Through a two-part study consisting of answering a questionnaire and responding to a post-questionnaire interview, six human subjects identified five cultural difference detection criteria: not only the (i) similarity of interpretations were considered, but also the (ii) percentage or ranking, (iii) conformity/variance, (iv) presence of proper nouns, and (v) positive/negative connotations in the interpretations were considered.

Consequently, some problems with using existing computational approaches to approximate human cultural difference detection were identified: the vector space model cannot distinguish structurally similar interpretation vector pairs; the hierarchical semantic net approach cancels distinctions between two interpretations or create artificial differences unfamiliar to human. Hence, cultural difference detection will require a new kind of computational approach in the future.

Chapter 5

Automatic Detection of Cultural Differences

This chapter describes a method of automatically detecting cultural differences in pictogram interpretations. Three human cultural difference detection criteria are formalized to define three inequalities, and these inequalities are connected using conjunctions to define a unified function which detects cultural differences. The function extends the semantic relevance measure defined in chapter 3 to handle pictogram interpretations of two cultures. The detection performance of the unified function is evaluated using a test data.

5.1 Introduction

Pictograms have clear pictorial similarities with some object [Marcus 03], and one who can recognize the object depicted in the pictogram can interpret the meaning associated with the object. Universal signs such as road signs, direction boards at the airports, and symbols of sports played in the Olympics are examples of familiar pictograms used nowadays which convey particular information to a wide range of audiences. On the other hand, special pictograms are developed for AAC (Augmentative and Alternative Communication) to assist people with severe communication difficulties [Maharaj 80] or to allow children from different countries to communicate to one another without using to a third language [Takasaki 06].

Existing pictogram communication systems such as Minspeak [Baker 82] and IconText [Beardon 95] use a fixed set of icons and system-defined sentence generation procedures to create pictogram messages. By contrast, a pictogram email system [Takasaki 06] uses an open set of pictograms where new pictograms are continuously added to the existing set of pictograms. The email system provides a two-dimensional canvas interface where a user can freely place one or more pictograms onto the canvas to create pictogram messages; no system-defined pictogram sentence generation procedure is imposed on the user. In this chapter, we consider system like [Takasaki 06] which uses an unfixed set of pictograms as candidates for conveying intended meaning, and so the selection of the most relevant pictogram becomes the sentence creation strategy.

Selecting the most relevant pictogram is not easy since pictorial symbols are not always universally interpretable. For example, the cow is a source of nourishment to westerners who drink milk and eat its meat, but it is an object of veneration to many people in India. Hence, a picture of cow could be interpreted quite differently by Protestants and Hindus [Kolers 69]. When pictograms having culturally different interpretations are used in intercultural setting, misunderstanding may arise between participants having different cultural backgrounds. One way to prevent such misunderstanding is to switch the pictogram selected by the sender to a more culturally appropriate pictogram suited to the receiver. To do this, we need to detect pictograms having culturally different interpretations.

This chapter proposes a method of automatically detecting cultural differences in pictogram interpretations by formalizing three human cultural difference detection criteria. The findings of chapter 4 dealing with human cultural difference detection criteria are utilized [Cho 09]. Three out of five human cultural difference detection criteria are formalized to define inequalities for detecting cultural differences in pictogram interpretations. The equation extends the semantic relevance measure of one country discussed in chapter 3 [Cho 08b, Cho 08a] to handle pictogram interpretations across two countries. Bilingual dictionaries and thesaurus are incorporated to automatically extract bilingual pairs of pictogram interpretation words of two countries (languages). The three inequalities are then merged using conjunctions to define a unified cultural difference detection function.

When detecting cultural differences, however, it would be more beneficial if reasons behind cultural differences are explicitly conveyed rather than simply presented in a yes or no fashion; understanding the content of cultural difference would allow users to understand each other and the other's culture better. Comprehensibility is important in any system [Sengers 99] but still not much supported. We aim to support it in the area of pictogram by formalizing how humans assess cultural differences in pictograms.

5.2 Human Detection of Cultural Differences Revisited

An online pictogram survey was conducted in the U.S. and Japan for 14 months to collect cross-cultural pictogram interpretations of 120 pictograms. The two countries were selected for their cultural distinctness and ease of data gathering and analysis [Hall 76, Hofstede 05]. These pictogram interpretations were then used as stimuli in the human cultural difference detection study [Cho 09].

5.2.1 Characteristics of Pictogram Interpretations

Analyzing the 120 U.S.-Japan pictogram interpretations revealed the following characteristics in the pictogram interpretations. We explain by giving examples in Table 5.1:

- Geographically and/or psychologically proximate real-world subject matter is identified (e.g. “Tokyo Tower”, “Tokyo”, “Eiffel Tower”, “Paris”)
- The position of the interpretation word in the concept hierarchy is varied (e.g. “radio tower”, “cell phone tower”, “electricity tower” are

subconcepts of “tower” while “building” is a superconcept of “tower”. By contrast, “Eiffel Tower” is an instance of “tower”.)

- Concrete object(s) depicted in the pictogram is associated with some abstract concept (e.g. “Tower” is associated with “travel”, “tour”, or “technology”.)
- Most pictograms have polysemous interpretations (e.g. “Eiffel Tower”, “tower”, “Paris”, “travel”, “oil”)

When each country’s pictogram interpretation words are considered as a set of concepts, the relationship between U.S.–Japan pictogram interpretation words can be classified into one of the four relations given below:

Disjoint Relation Almost no common concept exists between U.S. and Japan.

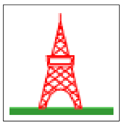
Intersecting Relation U.S. and Japan share some common concepts, but dissimilar concepts exist in both countries.

Subset Relation U.S. concepts subsume Japanese concepts or vice versa.

Equivalent Relation Most U.S. and Japanese concepts are the same.

It is natural to think that no cultural difference exists if most of the two countries’ interpretations are similar, i.e. the two countries’ concepts have an equivalent relation. The classification of relation, however, is dependent on the level of interpretations being handled: the interpretations can be handled either at the word-level or concept-level. For instance, U.S.-Japan interpretations in Table 5.1 can be classified as having an intersecting relation when viewed at the word-level since they share the word “tower”, but also have dissimilar interpretations such as “Eiffel Tower” and “Tokyo Tower”. By contrast, when viewed at the concept-level, both countries’ interpretations are centered on the concept of “tower”, so they can be classified as having an equivalent relation. In the former case, the two countries’ interpretations may be viewed as having cultural differences whereas in the latter case they can be viewed as having no cultural difference.

Table 5.1: Tallied U.S.–Japan pictogram interpretations



U.S.	FREQ	JAPAN	FREQ
Eiffel Tower	25	Tokyo Tower$_{exp1}$	41
Paris	25	tower$_{exp1}$	22
tower	20	tower$_{exp2}$	7
Eifel Tower	6	Eiffel Tower$_{exp1}$	6
France	6	radio tower$_{exp1}$	3
radio tower	3	TV tower	2
travel	3	Tokyo	2
building	2	Paris	1
oil rig	2	steel tower	1
cell phone tower	1	radio tower$_{exp2}$	1
electricity tower	1	Eiffel Tower$_{exp2}$	1
fun/expensive	1	tower$_{exp3}$	1
oil	1	Tokyo Tower$_{exp2}$	1
technology	1	Tokyo Tower$_{exp3}$	1
tour	1	tower$_{exp4}$	1
omitted words	32	tower (also Tokyo)	1
TOTAL	130	TOTAL	92

Note: Subscripts such as exp_1 denotes that Japanese words are expressed using different character combinations.

5.2.2 Human Cultural Detection Criteria

The U.S.–Japan pictogram interpretations were tallied as the one in Table 5.1 for each of the 120 pictograms. Three human subjects first identified sixty pictograms with possible cultural differences in 120 U.S.–Japan pictogram interpretations. Then the selected sixty pictogram interpretation tables were shown to different four human subjects; this time, the level of cultural differences was assessed using a seven-point Likert scale. The scale ranged from "Strong cultural difference exists (7.00)" to "Absolutely no cultural difference (1.00)". Thirty U.S.–Japan pictogram interpretations with greater cultural differences were then selected as stimuli for human cultural difference detection study.

Six new human subjects, three U.S. nationality and three Japanese, participated in a two-part study consisting of (a) answering a questionnaire and (b) responding to a post-questionnaire interview. Twelve out of thirty pictograms were identified as having some cultural differences (average of 5.00 or greater), and nineteen out of thirty were identified as having possible cultural differences (average of 4.00 or greater). The nineteen pictogram interpretations were assessed to have cultural differences based on the following criteria:

1. Few similar interpretations exist between the two countries.
2. Quite a few similar interpretations exist, but the percentages of those interpretations are different between the two countries.
3. Conformity of semantics is observed in one country's interpretations while variance is observed in the other.
4. Proper nouns such as the name of a country or city exist in the interpretations.
5. Negative connotation in the interpretations is observed in one country while positive connotation is observed in the other.

More than one criterion was mentioned as cultural difference detection criteria in each of the 19 pictogram interpretations with possible cultural differences.

5.3 Machine Detection of Cultural Differences

5.3.1 Formalization of Human Detection Criteria

We now formalize the human cultural difference detection criteria to define inequalities that automatically detect cultural differences in two countries' pictogram interpretations. The first three human detection criteria in section 5.2.2 deals with semantic similarity of U.S.–Japan interpretation words whereas the last two deals with proper nouns and positive/negative connotations. Since the detection of proper nouns can be handled by referring to a proper noun database, this criterion is excluded from the target of formalization. Also, judging positive/negative connotations needs some kind of an outside measure, this criterion is also excluded. Consequently, we will formalize the first three human detection criteria which deal with semantic similarity in the two countries interpretation words.

Before we formalize each of the three criteria, we first define a semantic similarity measure of the two interpretation words a_i, a_j in country A as follows:

$$Sim_A(a_i, a_j) = \frac{|Pict(a_i) \cap Pict(a_j)|}{|Pict(a_i) \cup Pict(a_j)|}$$

Here, $Pict(a_i)$, $Pict(a_j)$ each indicate a set of pictograms containing a_i, a_j as interpretation words respectively. For example, let's say there are 120 pictograms, and we wish to calculate the similarity of two words "happy" and "laughing"; we count the number of pictograms having "happy" (let's say there are 22), then we count the number of pictograms having "laughing" (let's say 5), and we count the number of pictograms having both "happy" and "laughing" (let's say 5). Then, the similarity value of "happy" and "laughing" $Sim_{US}(happy, laughing)$ is 5/22=0.227. The same similarity measure is defined for the two interpretation words in country B.

We can extend the similarity measure of one country to two countries A and B by defining the similarity measure of two countries' interpretation words a_i, b_j as follows:

$$\begin{aligned} &Sim_{A,B}(a_i, b_j) \\ &= \max_{a_k \in I_A, b_l \in I_B} \{d_{kl} Sim_A(a_i, a_k) Sim_B(b_j, b_l)\} \end{aligned}$$

Here, I_A, I_B each indicate a set of interpretation words in countries A and B for the same pictogram ($I_A = \{a_1, a_2, ..., a_m\}$, $I_B = \{b_1, b_2, ..., b_n\}$).

Furthermore, d_{kl} indicates whether the two countries' interpretation words a_k, b_l are bilingual pairs in a dictionary or related words.

$$d_{kl} = \begin{cases} 1 & (a_k, b_l : \ bilingual \ \ pairs \ \ or \ \ related) \\ 0 & (a_k, b_l : \ neither \ \ bilingual \ \ pairs \ \ nor \ \ related) \end{cases}$$

The value of d_{kl} is determined using a bilingual dictionaries and thesauri of countries A and B by looking up whether a_k, b_l are bilingual pairs or related words. For example, if the countries are U.S. and Japan, we can use a Japanese-English bilingual dictionary to look up whether Japanese interpretation word a_k and English interpretation word b_l are listed as bilingual pairs in the dictionary.

Now we are ready to formalize the three human cultural difference detection criteria. We formalize the first criterion as Rule 1 and define the following inequality:

Rule 1. There are few similar interpretations between two countries.

$$\sum_{i}^{m} \sum_{j}^{n} P(a_i \mid I_A) P(b_j \mid I_B) Sim_{A,B}(a_i, b_j) < thres_1 \tag{5.1}$$

$P(a_i \mid I_A)$ indicates the probability of interpretation word a_i in country A, and $P(b_j \mid I_B)$ indicates the probability of interpretation word b_j in country B. The left-hand expression takes the value between 0 and 1, and when the left-hand side value is less than $thres_1$, then cultural difference exists is returned. Note that the left-hand side expression extends the *semantic relevance measure* proposed by [Cho 08b] which defines semantic relevance between a query (interpretation word) and a set of pictogram interpretation words. We extend the semantic relevance measure that deals with one country's interpretations to handle two countries' interpretations; the left-hand side expression calculates the semantic relevance between the two sets of pictogram interpretations words of countries A and B. If the two countries' interpretation words' semantic relevance is low, than the calculated value will be low; if the value is lower than $thres_1$, then cultural difference exist is returned. Note that the initial threshold can be set by generating a list of arbitrary threshold (for example, a list of threshold with 0.01 intervals), and choosing the threshold which gives maximum precision or recall.

Next, we formalize the second cultural difference criterion as follows:

Rule 2. The topmost interpretation in country A does not exist in the country B, or exists but has a very low ratio, and vice versa is true for country B to A.

$$\sum_{i}^{m} P(a_i \mid I_A) Sim_{A,B}(a_i, b_1) < thres_{2a} \tag{5.2}$$

$$\wedge \sum_{j}^{n} P(b_j \mid I_B) Sim_{A,B}(a_1, b_j) < thres_{2b} \tag{5.3}$$

Note that each interpretation word in I_A, I_B are sorted according to the descending order of probability; that is, a_1, b_1 are interpretation words with largest probability in I_A, I_B. When the two inequalities hold true, then cultural difference exists is returned. What rule 2 is looking at is the major interpretation word of the two countries; when the two countries' major interpretations differ, than cultural difference exists is returned. Note that the initial thresholds ($thres_{2a}$, $thres_{2b}$) can be set by generating possible threshold pairs and choosing the pair with greatest detection performance.

Finally, we formalize the third cultural difference criterion as follows:

Rule 3. The interpretations in one country have high semantic conformity whereas the interpretations in another country have low semantic conformity.

First, we define the semantic conformity ($Conformity_A, Conformity_B$) of pictogram interpretation words in countries A and B as follows:

$$Conformity_A = \sum_{i}^{m} \sum_{j}^{m} P(a_i \mid I_A) P(a_j \mid I_A) Sim_A(a_i, a_j)$$

$$Conformity_B = \sum_{i}^{n} \sum_{j}^{n} P(b_i \mid I_B) P(b_j \mid I_B) Sim_B(b_i, b_j)$$

Each of the two equations calculates how much the interpretation words in one country are semantically related to one another. After each country's semantic conformity is calculated, the difference of the two countries' interpretation words' semantic conformity is calculated and the absolute value is compared to the threshold; if the absolute difference is greater than $thres_3$, than cultural difference exists is returned. Below is the inequality that defines rule 3:

$$| Conformity_A - Conformity_B |> thres_3 \tag{5.4}$$

Note that if the interpretation words in one country is semantically related to each other, the value of $Conformity$ will increase, but if they are varied, the $Conformity$ value will decrease. When the absolute difference of the two countries' $Conformity$ values are great, it means that the semantic conformities of the two countries' interpretation words are very different; and by this, cultural difference exists is returned. Note that the initial threshold $thres_3$ can be set in the same way as inequality (1), but the inequality sign is reversed.

5.3.2 Unified Function

The inequalities (1), (2), (3), (4) in section 5.3.1 are merged using conjunctions to define a unified function that automatically detects cultural differences in pictogram interpretations. The unified function is defined as follows:

$$\begin{aligned}\sum_{i}^{m}\sum_{j}^{n} P(a_i \mid I_A)P(b_j \mid I_B)Sim_{A,B}(a_i, b_j) < thres_1 \\ \wedge \sum_{i}^{m} P(a_i \mid I_A)Sim_{A,B}(a_i, b_1) < thres_{2a} \\ \wedge \sum_{j}^{n} P(b_j \mid I_B)Sim_{A,B}(a_1, b_j) < thres_{2b} \\ \wedge \mid Conformity_A - Conformity_B \mid > thres_3\end{aligned}$$

5.4 Experiment

An experiment was conducted to evaluate the detection performance of the three inequalities and the unified function using a 120 U.S.–Japan pictogram interpretation test data. For mapping Japanese and U.S. interpretation words (d_{kl}), two Japanese-English bilingual dictionaries “EDICT”[*] and “EIJIRO”[†], and EDR Concept Dictionary[‡], a Japanese thesaurus “Weblio” [§], and a list of English words in lemma and inflected forms[¶] were used to automatically extract Japanese and English bilingual or related word pairs.

5.4.1 Test Data

U.S.–Japan pictogram interpretation words of 120 pictograms were gathered from a web survey to generate a test data. 19 out of 120 pictogram interpretation words were designated as culturally different pictogram interpretations. The total U.S. and Japan respondents were 1,661 and 662 respectively. The average U.S. and Japanese respondents for each pictogram was 204 and 92 respectively. The U.S.–Japan interpretation words were cleaned and merged before being tallied: the fluctuation of expressions in Japanese words were merged based on lemma and pronunciation using a Japanese morphological analyzer ChaSen[‖]; misspellings and inflected forms in English words were merged based on edit distance and lemma. An example of tallied test data is shown in Table 5.2.

5.4.2 Sample Calculation

How to calculate the value of the left-hand expression of the inequality (1) (Rule 1) is explained using the data in Table 5.2. Seven and six interpretation

[*] http://www.csse.monash.edu.au/~jwb/j_edict.html
[†] http://www.alc.co.jp/ CR-ROM version, 3rd edition.
[‡] http://www2.nict.go.jp/r/r312/EDR/index.html
[§] http://thesaurus.weblio.jp/
[¶] http://www.lexically.net/downloads/e_lemma.zip
[‖] http://chasen.naist.jp/hiki/ChaSen/

words are given as U.S. and Japan interpretations respectively; we calculate all the values for the forty-two U.S.–Japan pairs ($7 \times 6 = 42$). For instance, let's say we want to calculate the value of the topmost "Eiffel Tower$_{U.S.}$" and "Tokyo Tower": $P(a_i \mid I_A)$ is 0.443, $P(b_j \mid I_B)$ is 0.485, and we are to calculate the product of these two values and $Sim_{A,B}(a_i, b_j)$. To calculate the similarity value $Sim_{A,B}(a_i, b_j)$, we use bilingual dictionaries to extract bilingual pairs that will return the greatest $Sim_{A,B}(a_i, b_j)$ value. In the case of "Eiffel Tower$_{U.S.}$" and "Tokyo Tower", the only bilingual pair available in the dictionary is "Eiffel Tower$_{U.S.}$" and "Eiffel Tower$_{Japan}$", so we indirectly calculate the similarity of "Eiffel Tower$_{Japan}$" and "Tokyo Tower" by obtaining the product of the following three values:

- $d_{kl} = 1$: "Eiffel Tower$_{U.S.}$", "Eiffel Tower$_{Japan}$"
- $Sim_A(a_i, a_k) = 1.0$: similarity value between "Eiffel Tower$_{U.S.}$" and "Eiffel Tower$_{U.S.}$"
- $Sim_B(b_j, b_l) = 0.5$: similarity value between "Eiffel Tower$_{Japan}$" and "Tokyo Tower"

All forty-two U.S.–Japan interpretation word pairs are calculated similarly, and the overall sum is obtained as the semantic relevance value for the U.S.–Japan interpretations of the given pictogram. The actual value for the Table 5.2 data is 0.65922. If this value is smaller than $thres_1$, then cultural difference exists is returned.

The left-hand expressions for inequalities (2) and (3) (Rule 2) can be calculated similarly, but the inputs are the topmost interpretation word of one country and all interpretations words of the other country. In the case of inequality (4) (Rule 3), similar calculation is applied but the inputs are the two sets of the same interpretation words of one country; no bilingual dictionary is used for mapping two languages since the two sets of data are from the same country (same language).

5.4.3 Setting Threshold

We generated a set of arbitrary threshold values for inequalities (1) and (4) (Rules 1 and 3) with 0.02 intervals, and chose the threshold with the highest *F*-measure. For inequalities (2) and (3) (Rule 2), a pair of left-hand expression values for 19 culturally different pictograms were used as thresholds ($thres_{2a}$, $thres_{2b}$), and the pair with the greatest *F*-measure was chosen. The precision, recall, *F*-measure, and threshold of each inequalities (Rules R1, R2, and R3) are organized in Table 5.4.

One out of 120 pictogram interpretation data were retrieved as test data for the unified function, and the remaining 119 data were used as training data; the same data generation was done for all 120 data (leave-one-out cross-validation). The four thresholds of the unified function (R1 $\wedge$ R2 $\wedge$ R3) were picked which returned the greatest *F*-measure. The single test data was then tested for cultural difference detection using the unified equation with the thresholds.

Of the 120 pictogram interpretation data, 19 data were culturally different. 105 out of 120 test data were correctly detected or classified for cultural differences (105/120=87.5%). Fifteen test data were incorrectly classified; 11 out of 15 were culturally different pictogram interpretations which had failed to be detected; 4 out of 15 were false positives (no cultural differences, but detected). The fifteen incorrectly detected cases were excluded, and the correctly detected 105 cases were used to set the thresholds of the unified function by averaging the 105 sets of thresholds.

Table 5.2: U.S.–Japan pictogram interpretation word test data used for detection performance evaluation

U.S.	%	Japan	%
Pictogram P28 in Chapter 4 Figure 4.2 (Table 5.3 Pict. No. 111)			
Eiffel Tower$_{U.S.}$	44.3	Tokyo Tower (東京タワー)	48.5
Paris	24.6	tower (タワー)	28.7
tower	21.2	tower (塔)	11.9
France	4.9	Eiffel Tower (エッフェル塔)	6.9
pairs	2.0	TV tower (テレビ塔)	2.0
building	1.5	Tokyo (東京)	2.0
travel	1.5		
Pictogram P04 in Chapter 4 Figure 4.2 (Table 5.3 Pict. No. 027)			
man	41.1	man (男の人)	38.8
dad	16.4	man (男性)	31.1
woman	10.3	man (男)	15.5
adult	7.9	father (お父さん)	9.7
father	7.5	adult (大人)	4.9
mom	7.0		
Pictogram P01 in Chapter 4 Figure 4.2 (Table 5.3 Pict. No. 004)			
exercise	36.7	OK	27.8
stretch	11.6	circle (まる)	26.4
circle	8.2	correct (正解)	22.2
happy	7.5	OK (オッケー)	15.3
yoga	6.8	good (良い)	2.8
jump rope	6.1	OK (オーケー)	2.8
dance	2.7	meet (合う)	2.8
Pictogram P23 in Chapter 4 Figure 4.2 (Table 5.3 Pict. No. 089)			
winner	44.4	atheletic meet (運動会)	56.6
winning	9.5	victory (ゆうしょう)	14.5
champion	5.8	first place (一番)	13.2
first place	5.8	first place (一位)	6.6
race	3.7	win (勝つ)	3.9
cheer	3.3	compete (競争する)	2.6
win	2.9	rejoice (喜ぶ)	2.6

Table 5.3: Comparison of detection results of 19 culturally different pictogram interpretations using three inequalities (R1, R2, R3) and a unified function (R1 ∧ R2 ∧ R3)

Pict. No.	Independent			Unified	Human Judged	
	R1	R2	R3	R1 ∧ R2 ∧ R3	AVG	SD
086	✓	✓	✓	✓	7.00	0.00
004	✓	✓	✓	✓	6.83	0.37
077	✓	✓	✓	✓	6.33	0.75
111	✓	✓			6.17	0.37
005	✓	✓	✓	✓	6.17	0.69
080	✓	✓	✓	✓	6.00	0.58
076	✓	✓	✓	✓	6.00	1.41
078	✓	✓	✓	✓	5.83	1.46
083		✓	✓	✓	5.67	1.11
072	✓	✓	✓	✓	5.50	1.71
050	✓	✓		✓	5.17	1.07
114		✓		✓	5.00	0.58
087		✓	✓	✓	4.67	0.75
068		✓			4.67	1.25
069	✓	✓		✓	4.67	1.25
089					4.50	0.76
067	✓	✓			4.17	1.77
026		✓			4.00	1.41
027					4.00	1.41
Correct	12	17	10	13		
Wrong	15	22	8	5		
Total	27	39	18	18		

Table 5.4: Precision, recall, F_1-measure, and threshold

	R1	R2	R3	R1 $\wedge$ R2 $\wedge$ R3		
Precision	48.1	43.5	55.5	72.2%		
Recall	63.1	89.4	52.6	68.4%		
F_1-measure	52.1	58.6	54.0	70.2%		
threshold	0.66	0.401 0.365	0.36	0.78380	0.40234 0.36967	0.12790

5.4.4 Evaluation

Table 5.3 shows the detection result for 19 culturally different pictogram interpretations using the three detection inequalities (R1, R2, and R3) and using the unified function. The precision, recall, F_1-measure and thresholds are organized is Table 5.4. The precisions of each detection inequalities R1, R2, and R3 were 48.1% (12/27), 43.5% (17/39), and 55.5% (10/18) for respectively. The recalls were 63.1% (12/19), 89.4% (17/19), and 52.6% (10/19) respectively. Meanwhile, the unified fucntion's precision and recall were 72.2% (13/18) and 68.4% (13/19) respectively.

We may have used disjunctions ($\vee$) instead of conjunctions ($\wedge$) to define the unified function, but the former has a higher probability of returning a greater number of false positives; more importantly, the conjunctions approach coincides with the human cultural difference detection approach in chapter 4 ([Cho 09]) which describes that pictogram interpretations were assessed to have cultural differences based on two or more detection criteria.

5.5 Summary

We proposed a method of detecting cultural differences in two countries' pictogram interpretations by formalizing the three human cultural difference detection criteria on pictogram interpretations. Each detection criteria was defined as the inequality that calculates the semantic relevance value of the two countries' pictogram interpretations, or the absolute value difference of the two countries' internal semantic relevance values. The left-hand expression of each inequality indirectly utilized the similarity value of one country. Bilingual dictionaries and thesaurus were incorporated to automatically extract bilingual pairs of the two countries' interpretation words during calculation. The three inequalities (R1, R2, and R3) were merged using conjunctions to define a unified function that automatically detects cultural differences in pictogram interpretations. The precision and recall of the unified function on 120 pictogram interpretations test data containing 19 culturally different interpretations was 72.2% and 68.4% respectively.

Chapter 6

Conclusion

6.1 Contributions

Our goal was to propose a system-level handling and detection of cultural differences in human-provided cross-cultural pictogram interpretations. As an ultimate goal, we aimed to build an agent which can automatically detect cultural differences. Existing researches on culturally-situated agents have tackled the problem of cooperation between agents with different cultural backgrounds [Chaudron 98] or the problem of bridging humans with different cultural backgrounds [Ishida 06a]. The former focuses on conflict resolution while the latter focuses on mediation. In this work, we tackled the problem of automatically detecting cultural differences based on human-provided interpretations. We used pictogram as a symbolic medium to collect human interpretations from two different cultures. We tackled three issues to realize our goal:

- **Semantic relevance measure was proposed to tackle semantic ambiguity in pictogram interpretations.**
  Polysemous and shared pictogram interpretation can lead to ambiguity in pictogram interpretation, which can cause misunderstanding in communication using pictograms. To retrieve pictograms that can better convey the intended meaning, we proposed a method of select-

ing and ranking relevant pictograms which are more likely to be interpreted as intended. We proposed a categorical semantic relevance measure, which calculates how relevant a pictogram is to a given interpretation in terms of a pictogram category. The measure defines the probability and similarity measurement of categorized pictogram interpretations. Five pictogram categories used for categorizing pictogram interpretation words were defined using the Concept Dictionary of the EDR Electronic Dictionary. Three semantic relevance approaches, (i) not-categorized semantic relevance approach, (ii) categorized approach, and (iii) categorized and weighted approach, were evaluated, and the categorized approaches showed more stable performance than the not-categorized approach.

- **Human cultural difference detection criteria were clarified.**
  To understand whether human detection of cultural differences can be approximated using the existing computational approaches (for example, semantic similarity calculation), a study on how humans detect cultural differences in pictogram interpretations was conducted. Through a two-part study consisting of answering a questionnaire and responding to a post-questionnaire interview, six human subjects identified five cultural difference detection criteria: not only the (i) similarity of interpretations were considered, but also the (ii) percentage or ranking, (iii) conformity/variance, (iv) presence of proper nouns, and (v) positive/negative connotations in the interpretations were considered. Consequently, some problems with using existing computational approaches to approximate human cultural difference detection were identified: the vector space model cannot distinguish structurally similar interpretation vector pairs; the hierarchical semantic net approach cancels distinctions between two interpretations or creates artificial differences unfamiliar to human. Hence, cultural difference detection requires a new kind of computational approach.

- **Automatic method of detecting cultural differences in cross-cultural pictogram interpretations was proposed.**
  We proposed a method of detecting cultural differences in two countries' pictogram interpretations by formalizing the three human cultural difference detection criteria on pictogram interpretations. Each detection criteria was defined as an inequality that checks (i) the difference of the semantic relevance value between the two countries' overall pictogram interpretations, (ii) the difference of the semantic relevance between the two countries' majority pictogram interpretations, and (iii) the difference of semantic variance/conformity between the two countries' overall pictogram interpretations. Each inequality utilized the semantic relevance measure which was extended to handle pictogram interpretations of two cultures. Bilingual dictionaries and thesaurus were incorporated to automatically extract bilingual pairs of the two countries' interpretation words during calculation. The three inequalities were merged using conjunctions to define a unified function that automatically detects cultural differences in pictogram interpretations. The detection performance of the unified function on U.S.–Japan pictogram interpretation test data was approximately 70% (F_1-measure in information retrieval).

6.2 Future Directions

Imagine a shop attendant agent which suggests to the human customer a red color gift-wrapping over blue one after checking the gift receiver's name to be a typical Chinese name (red is considered auspicious to many Chinese), or a chef agent which outputs a dinner menu with hot spicy main dish after being informed that human guests coming to dinner are two Koreans and a Thai. Agents which can detect cultural differences are useful in many respects and can have numerous application areas. For example, we can build a system which collects cross-cultural interpretations by showing various images, photos, and illustrations and detect cultural differences in those interpretations to identify what kinds of graphic elements are independent, neutral, or universal regardless of culture and what elements are culture-specific or sensitive to culture. We plan to seek real-world application area where the automatic detection of cultural differences in human-provided interpretations can be useful and implement novel culture-sensitive applications in the future.

Bibliography

[Aurnhammer 06] Aurnhammer, M., Hanappe, P., and Steels, L.: Augmenting navigation for collaborative tagging with emergent semantics, in *ISWC '06 Proc. of the 5th Int'l Semantic Web Conf.*, Vol. 4273 of *LNCS*, pp. 58–71, Springer (2006)

[Baker 82] Baker, B. R.: Minspeak, A semantic compaction system that makes self-expression easier for communicatively disabled individuals, *Byte*, Vol. 7, No. 9, pp. 186–202 (1982)

[Beardon 95] Beardon, C.: Discourse Structures in Iconic Communication, *Artificial Intelligence Review*, Vol. 9, pp. 189–203 (1995)

[Bliss 65] Bliss, C. K.: *Semantography (Blissymbolics)*, Semantography (Blissymbolics) Publications (1965)

[Champoux 00] Champoux, B., Inoue, T., Iwadate, Y., and Fujisawa, K.: Transmitting Visual Information: Icons Become Words, in *Proceedings of the IEEE International Conference on Information Visualisation (IV-00)*, pp. 244–249, IEEE (2000)

[Chaudron 98] Chaudron, L., Erceau, J., and Tessier, C.: Cultural differences? An opportunity for a mutual knowledge enhancement, in *PRICAI'98 Workshop on 'Issues in cross-cultural communication: Toward culturally situated agents'* (1998)

[Cho 07a] Cho, H., Ishida, T., Inaba, R., Takasaki, T., and Mori, Y.: Pictogram Retrieval Based on Collective Semantics, in Jacko, J. ed., *Human-*

Computer Interaction. HCI Intelligent Multimodal Interaction Environments, Vol. 4552 of *LNCS*, pp. 31–39, Springer (2007)

[Cho 07b] Cho, H., Ishida, T., Yamashita, N., Inaba, R., Mori, Y., and Koda, T.: Culturally-Situated Pictogram Retrieval, in Ishida, T., Fussell, S. R., and Vossen, P. eds., *Intercultural Collaboration*, Vol. 4568 of *LNCS*, pp. 221–235, Springer (2007)

[Cho 08a] Cho, H., Ishida, T., Oyama, S., Inaba, R., and Takasaki, T.: Assisting Pictogram Selection with Categorized Semantics, *IEICE Transactions on Information and Systems*, Vol. E91-D, No. 11, pp. 2638–2646 (2008)

[Cho 08b] Cho, H., Ishida, T., Takasaki, T., and Oyama, S.: Assisting Pictogram Selection with Semantic Interpretation, in *ESWC '08 Proc. 5th European Semantic Web Conf.*, Vol. 65–79 of *5021*, p. LNCS, Springer (2008)

[Cho 09] Cho, H., Ishida, T., Yamashita, N., Koda, T., and Takasaki, T.: Human Detection of Cultural Differences in Pictogram Interpretations, in *Proc. 2nd International Workshop on Intercultural Collaboration*, ACM Press (2009)

[EDR 02] EDR, E.: *EDR Electronic Dictionary Version 2.0 Technical Guide*, National Institute of Information and Communications Technology (NICT) (2002), Available at: `http://www2.nict.go.jp/r/r312/EDR/`

[Ekman 03] Ekman, P.: *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*, Henry Holt & Co. (2003)

[Geertz 73] Geertz, C.: *The Interpretation of Cultures*, Basic Books (1973)

[Hall 76] Hall, E.: *Beyond Culture*, Doubleday & Company (1976)

[Hofstede 05] Hofstede, G. and Hofstede, G. J.: *Cultures and Organizations: Software of the mind*, McGraw Hill (2005)

[Ishida 06a] Ishida, T.: Communicating Culture, *IEEE Intelligent Systems, Special Issue on the Future of AI*, Vol. 21, No. 3, pp. 62–63 (2006)

[Ishida 06b] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, in *SAINT '06 Proc. Int'l Symposium on Applications and the Internet*, pp. 96–100, IEEE Computer Society (2006)

[Koda 06] Koda, T. and Ishida, T.: Cross-Cultural Study of Avatar Expression Interpretations, in *SAINT '06 Proc. Int'l Symposium on Applications and the Internet*, pp. 130–136, IEEE Computer Society (2006)

[Kolers 69] Kolers, P. A.: Some Formal Characteristics of Pictograms, *American Scientist*, Vol. 57, pp. 348–363 (1969)

[Kroeber 52] Kroeber, L. and Kluckhohn, C.: Culture: A Critical Review of Concepts and Definitions, *Harvard University Peabody Museum of American Archaeology and Ethnology Papers*, Vol. 47, (1952)

[Lin 98] Lin, D.: An information-theoretic definition of similarity, in *ICML '98 Proc. of the 15th Int'l Conf. on Machine Learning*, pp. 296–304 (1998)

[Maharaj 80] Maharaj, S. C.: *Pictogram Ideogram Communication*, Saskatoon, SK: The Pictogram Centre, Saskatchewan Association of Rehabilitation Centres (1980)

[Manning 99] Manning, C. D. and Schütze, H.: *Foundations of Statistical Natural Language Processing*, MIT Press (1999)

[Marcus 03] Marcus, A.: Icons, Symbols, and Signs: Visible Languages to Facilitate Communication, *Interactions*, Vol. 10, No. 3, pp. 37–43 (2003)

[Marcus 07] Marcus, A.: m-LoCoS UI: A Universal Visible Language for Global Mobile Communication, in Jacko, J. ed., *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, Vol. 4552 of *LNCS*, pp. 144–153, Springer (2007)

[Mori 07] Mori, Y.: Atoms of Bonding: Communication Components Bridging Children Worldwide, in Ishida, T., Fussell, S. R., and Vossen, P. eds., *Intercultural Collaboration*, Vol. 4568 of *LNCS*, pp. 335–343, Springer (2007)

[Morris 38] Morris, C. W.: Foundations of the Theory of Signs, in *International Encyclopedia of Unified Science*, Vol. 1, The University of Chicago Press (1938)

[Niles 01] Niles, I. and Pease, A.: Towards a standard upper ontology, in *FOIS '01 Proc. 2nd Int'l Conf. on Formal Ontology in Information Systems* (2001)

[Nisbett 01] Nisbett, R. E., Peng, K., Choi, I., and Norenzayan, A.: Culture and Systems of Thought: Holistic Versus Analytic Cognition, *Psychological Review*, Vol. 108, No. 2, pp. 291–310 (2001)

[Olson 03] Olson, J. S. and Olson, G. M.: Culture Surprises in Remote Software Development Teams, *ACM Queue*, Vol. 1, No. 9, pp. 52–59 (2003)

[Osgood 57] Osgood, C. E., Suci, G. J., and Tannenbaum, P. H.: *The Measurement of Meaning*, The University of Illinois Press (1957)

[Osgood 75] Osgood, C. E., May, W. H., and Miron, M. S.: *Cross-cultural Universals of Affective Meaning*, The University of Illinois Press (1975)

[Palmer 05] Palmer, M., Gildea, D., and Kingsbur, P.: The proposition bank: An annotated corpus of semantic roles, *Computational Linguistics*, Vol. 31, No. 1, pp. 71–105 (2005)

[Rada 89] Rada, R., Mili, H., Bicknell, E., and Blettner, M.: Development and applications of a metric on semantic nets, *IEEE Transactions on System, Man and Cybernetics*, Vol. 19, No. 1, pp. 17–30 (1989)

[Rijsbergen 79] Rijsbergen, van C.: *Information Retrieval*, Butterworths (1979)

[Romney 97] Romney, A. K., Moore, C. C., and Rusch, C. D.: Cultural universals: Measuring the semantic structure of emotion terms in English and Japanese, *PNAS: Proc. Natl. Acad. Sci. USA*, Vol. 94, pp. 5489–5494 (1997)

[Sengers 99] Sengers, P.: Designing Comprehensible Agents, in *IJCAI '99 Proc. 16th Int'l Joint Conf. on Artificial Intelligence*, pp. 1227–1232, Morgan Kaufmann Publishers Inc. (1999)

[Setlock 04] Setlock, L. D., Fussell, S. R., and Neuwirth, C.: Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging, in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pp. 604–613, ACM (2004)

[Takasaki 06] Takasaki, T.: PictNet: Semantic Infrastructure for Pictogram Communication, in Sojka, P., Choi, K.-S., Fellbaum, C., and Vossen, P. eds., *GWC '06 Proc. 3rd Int'l WordNet Conf.*, pp. 279–284, Masaryk University in Brno (2006)

[Takasaki 07] Takasaki, T.: Design and Development of a Pictogram Communication System for Children around the World, in Ishida, T., Fussell, S. R., and Vossen, P. eds., *Intercultural Collaboration*, Vol. 4568 of *LNCS*, pp. 193–206, Springer (2007)

[Uchimoto 05] Uchimoto, K., Hayashida, N., Ishida, T., and Isahara, H.: Automatic Rating of Machine Translatability, in *MT Summit X: Proc. 10th Machine Translation Summit*, pp. 235–242 (2005)

[Vanhauer 07] Vanhauer, M., Oertel, K., and Vöskamp, J.: Overcoming the Language Barrier: The Potential of the Visual Language LoCoS in International Human-Computer Communication, in Aykin, N. ed., *Usability and Internationalization, Part II, HCII 2007*, Vol. 4560 of *LNCS*, pp. 527–536, Springer (2007)

[Vatrapu 07] Vatrapu, R. and Suthers, D.: Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning, in Ishida, T., Fussell, S. R., and Vossen, P. eds., *Intercultural Collaboration*, Vol. 4568 of *LNCS*, pp. 260–275, Springer (2007)

[Yamashita 06] Yamashita, N. and Ishida, T.: Automatic prediction of misconceptions in multilingual computer-mediated communication, in *IUI '06 Proc. 11th Int'l Conf. on Intelligent User Interfaces*, pp. 62–69, ACM Press (2006)

Publications

Major Publications

Journals

1. Heeryon Cho, Toru Ishida, Satoshi Oyama, Rieko Inaba, and Toshiyuki Takasaki, "Assisting Pictogram Selection with Categorized Semantics," Special Section on Knowledge, Information and Creativity Support System, *IEICE Transactions on Information and Systems*, Vol. E91-D, No. 11, pp. 2638–2646, 2008.

2. Heeryon Cho, Toru Ishida, Naomi Yamashita, Rieko Inaba, Toshiyuki Takasaki, and Tomoko Koda, "Human Judgment of Cultural Differences in Pictogram Interpretations," *Journal of Human Interface Society*, Vol. 10, No. 4, pp. 427–434, 2008 (in Japanese).

International Conference

1. Heeryon Cho, Toru Ishida, Toshiyuki Takasaki, and Satoshi Oyama, "Assisting Pictogram Selection with Semantic Interpretation," *European Semantic Web Conference (ESWC-08)*, Lecture Notes in Computer Science, 5021, Springer-Verlag, pp. 65–79, 2008.

2. Heeryon Cho, Naomi Yamashita, and Toru Ishida, "Towards Culturally-Situated Agent Which Can Detect Cultural Differences," *Pacific Rim International Conference on Multi-Agents (PRIMA-07)*,

short paper, Lecture Notes in Artificial Intelligence, 5044, Springer-Verlag, pp. 458–463, 2008.

3. Heeryon Cho, Toru Ishida, Rieko Inaba, Toshiyuki Takasaki, and Yumiko Mori, “Pictogram Retrieval Based on Collective Semantics,” *International Conference on Human-Computer Interaction (HCII-07)*, Lecture Notes in Computer Science, 4552, Springer-Verlag, pp. 31–39, 2007.

Other Publications

Workshops

1. Heeryon Cho, Toru Ishida, Naomi Yamashita, Tomoko Koda, and Toshiyuki Takasaki, “Human Detection of Cultural Differences in Pictogram Interpretations,” *International Workshop on Intercultural Collaboration (IWIC-09)*, February 20-21, 2009, Stanford, CA, U.S.A.

2. Heeryon Cho, Toru Ishida, Naomi Yamashita, Rieko Inaba, Yumiko Mori, and Tomoko Koda, “Culturally-Situated Pictogram Retrieval,” *International Workshop on Intercultural Collaboration (IWIC-07)*, Lecture Notes in Computer Science, 4568, Springer-Verlag, pp. 221–235, 2007.

3. Shogo Kamiya, Tomoko Koda, Toshiyuki Takasaki, and Heeryon Cho, “Development of Pictogram Conversion Tool for Cross-culture Collaboration,” *Human-Agent Interaction Symposium (HAI-07)*, December 5–6, 2007 (in Japanese).

4. Heeryon Cho, Rieko Inaba, Toru Ishida, Toshiyuki Takasaki, and Yumiko Mori, “Semantics in Pictogram Communication,” *Joint Agent Workshops and Symposium (JAWS-06)*, October 25–27, 2006 (in Japanese).

Conventions

1. Heeryon Cho, Toru Ishida, Rieko Inaba, Toshiyuki Takasaki, and Yumiko Mori, "Semantic Relevance Measure Using Pictogram Interpretation Words," *IEICE National Convention*, March 20–23, 2007 (in Japanese).