Emotion Recognition from Text Using Semantic Labels and Separable Mixture Models

CHUNG-HSIEN WU, ZE-JING CHUANG, AND YU-CHUNG LIN
National Cheng Kung University, Tainan, Taiwan

This study presents a novel approach to automatic emotion recognition from text. First, emotion generation rules (EGRs) are manually deduced from psychology to represent the conditions for generating emotion. Based on the EGRs, the emotional state of each sentence can be represented as a sequence of semantic labels (SLs) and attributes (ATTs); SLs are defined as the domain-independent features, while ATTs are domain-dependent. The emotion association rules (EARs) represented by SLs and ATTs for each emotion are automatically derived from the sentences in an emotional text corpus using the *a priori* algorithm. Finally, a separable mixture model (SMM) is adopted to estimate the similarity between an input sentence and the EARs of each emotional state. Since some features defined in this approach are domain-dependent, a dialog system focusing on the students' daily expressions is constructed, and only three emotional states, *happy, unhappy, and neutral,* are considered for performance evaluation. According to the results of the experiments, given the domain corpus, the proposed approach is promising, and easily ported into other domains.

Categories and Subject Descriptors: I.2.7 [**Artificial Intelligence**]: Natural Language Processing -- *Text analysis*; I5.4 [**Pattern Recognition**]: Applications --*Text processing*

General Terms: Languages, Performance
Additional Key Words and Phrases: Emotion extraction

1. INTRODUCTION

Recent research into human-machine communication places greater emphasis on recognition of nonverbal information, especially of emotional reactions. An extensive literature has shown that it is helpful for communication systems to identify the users' emotional states [Picard et al. 2001]. The recognition of emotion is an important component of "affective computing" and has been implemented in many kinds of media. For example, speech and image are the most common ways to recognize emotion. Picard [1997] proposed the concept of affective computing and deduced theories of emotion, recognition, and generation. Cowie et al. [2001] initiated research on recognizing emotion via all kinds of signals, including emotional keywords, speech signals, and facial expressions. Some researchers proposed a multimodal system for emotion recognition: For example, Cohn and Katz [1998] developed an automated method for recognizing emotions via facial expressions and speech intonations. De Silva and Ng [2000] used a hidden Markov model (HMM) to recognize emotions from both video and audio signals.

This research was supported by the Ministry of Economic Affairs of the Republic of China under contract 92-EC-17-A-02-S1-024.
Authors' addresses: C.H. Wu, Z.J. Chuang, and Y.C. Lin, Dept. of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan, 701; email:{chwu, bala}@csie.nckue.du.tw, p7690106@dec4000.cc.ncku.edu.tw

Among the above approaches, recognizing emotion from speech is the most popular. In speech-based recognition, paralinguistic information such as pitch and energy are the most important acoustic features [Yu et al. 2001; Amir et al. 2003]; other significant features include zero-crossing-rate (ZCR) [Kwon et al. 2003]; Teager energy operator-based parameters (TEO-based parameters) [Rahurkar and Hansen 2003]; and frequency-centroid, formant, and pitch vibration.

Although communication systems can identify the users' emotional states from different communication modalities, the variety and complexity of language makes it difficult for researchers to recognize emotional states from pure textual data. Recognizing emotion is extremely important for some text-based communication tools, e.g., the dialog system is a kind of human machine communication system that uses only text input and output. Recognizing the users' emotional states enables the dialog system to change the response and answer types [Lee et al. 2002]. Text is still the main communication tool on the Internet. In online chat and BBS systems, the users' emotional states can be used to control the dialog strategy. In net conferencing, the ability to identify the users' emotional states can reduce the translation data size and further increase the fluency of the conferencing program [Boucouvalas 2002].

This research method recognizes the emotional state from textual input; the motivation for this research is twofold. First, textual data is the most popular medium, consisting of books, newspapers, and letters. And due to its small storage requirements, textual data is the most appropriate medium for network transmissions. Second, the variety and complexity of textual data makes it possible for people to exchange ideas, opinions, and emotions using text only.

Traditionally, research on the recognition of emotion from text was focused on the discovery and utilization of emotional keywords, that is, specific words that express the speaker's emotional state. Using emotional keywords is the most direct way to recognize a user's emotions from text input, and several methods were proposed that used selected emotional keywords. Yanaru [1995] took footage of real speakers while they were talking using emotional keywords and defined the emotional states corresponding to the keywords. Subasic and Huettner [2001] classified a group of emotional words by manually scoring the emotion level for each word. Boucouvalas and Zhe [2002] applied a parser to identify the objects associated with the emotional keywords. Devillers et al. [2002; 2003] found the most appropriate emotional state by calculating the conditional probability between the emotional keywords and the emotional states. Tao and Tan [2004] used emotional function words instead of emotional keywords to evaluate emotional states. However, all keyword-based systems have the following problems: (1) ambiguity in defining all emotional keywords; (2) recognizing emotion from sentences with no emotional keywords; and most importantly (3) lack of semantic and syntactic information for emotion recognition. Besides keyword-based approaches, some researchers utilized other textual information clues such as pragmatic intent, text content plausibility, and paragraph structure [Dijkstra et al. 1994]. Litman and Forbes [2003] integrated all dialog information, including acoustic and linguistic features, dialog acts, and the sequence of speakers, in order to recognize the emotional state in a dialog system [Forbes-Riley and Litman 2004]. Schuller et al [2004] integrated both acoustic and linguistic information in emotion recognition. These methods do not take the semantic and syntactic information in the textual data into account, and can only be used for an article with a complete chapter structure. With further analysis, some researchers believe that textual data is rich with emotion at the semantic level; that is, that emotional content is also contained in the semantic structure. Chan and Franklin [1998] analyzed input

sentences and constructed a symbolic network to reduce language model perplexity. Woods [1970] used a transition network to analyze natural language. A semantic network-based emotion recognition mechanism [Chuang and Wu 2002] was proposed using emotional keywords, semantic/syntactic information, and emotional history in order to recognize the emotional state of a speaker.

In this article we propose an emotion-recognition approach that uses textual content -- a diagram of this approach is shown in Figure 1. There are two main phases in the diagram: a training phase and a test phase. The training phase receives the emotion generation rules (EGRs) and emotional corpus to generate emotion association rules (EARs). The separable mixture models (SMMs) are applied to classify the most appropriate emotional state from the input text. In both phases a universal lexical ontology is employed. An exhaustive description of the procedures is shown in the following sections.

The rest of this article is organized as follows. An overview of the emotion recognition procedure is described in the next section. The details of each component are introduced in Sections 3 to 6. Section 3 describes the derivation of the emotion generation rules. Section 4 describes the definition of SLs. Section 5 introduce the *a priori* algorithm and the process for generating EARs. The separable mixture model is described in Section 6. Sections 7 and 8 show the experimental results and conclusions, respectively.

2. OVERVIEW OF EMOTION RECOGNITION PROCEDURES

An overview of our proposed approach is shown in Figure 1. In the training phase, the emotion generation rules (EGRs) are deduced manually to describe the conditions for generating emotion. All of the sentences in the collected emotional corpus are then annotated with the emotional state and EGR to which they belong. With further analysis, each EGR is seen to consist of domain-independent and domain-dependent components.

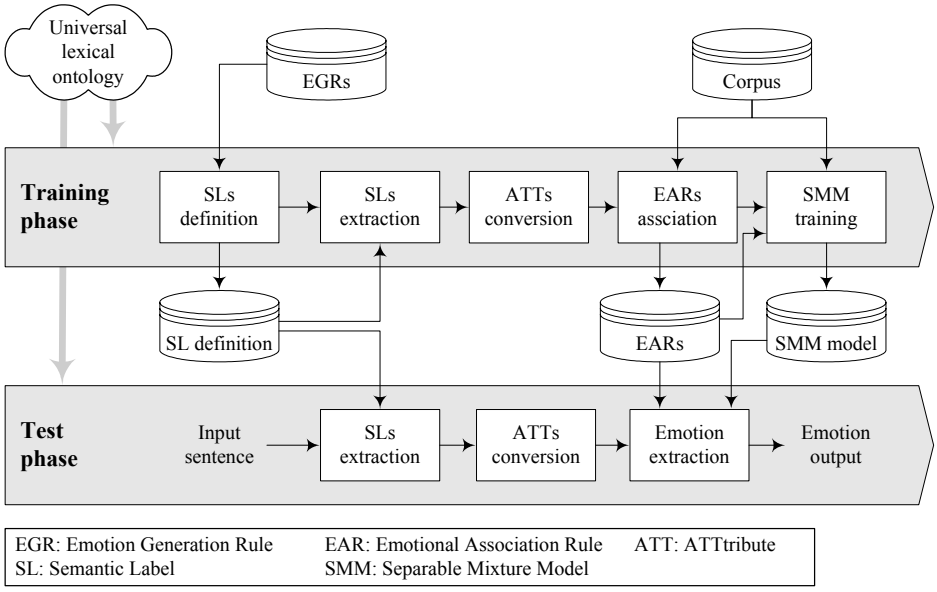

Fig.1. System block diagram.

For example, one EGR for the emotional state "HAPPY" is

"*One may be HAPPY if he obtains something beneficial*."

The domain-independent component is the semantic item "obtains," while the domain- dependent component is the semantic item "something that is beneficial." A sentence with this EGR annotation implies that we can extract these two components from the sentence. To obtain the domain-independent component, we define semantic labels (SLs) by analyzing the hierarchical hypernym structure of "obtains," which is already defined in a universal lexical ontology. Every word that contains predefined hypernyms can be converted into the corresponding SL. In the opposite direction, the domain-dependent components cannot be extracted as directly because they have no general definition. With the same example, the definition of "what is beneficial" is different in different domains and for different persons. Our approach to extracting these domain-dependent components is to train emotion association rules (EARs) using the *a priori* algorithm [Han and Kamber 2001]. Before the training process, a preprocessor converts all sentences in the training corpus into a sequence of SLs, ATTributes (ATTs) (defined in the universal lexical ontology mentioned above), and their inferential emotional state. The format of the converted result is

$$SL_1 + \ldots + SL_m + ATT_1 + \ldots + ATT_n \rightarrow [\text{EMOTIONAL\_STATE}].$$

Then, by applying the *a priori* algorithm, the so-called EARs consisting of SLs, the ATTs and the inferential emotional state are extracted. Because the main idea of emotion recognition is to evaluate the similarity between an input sentence and the EARs of each emotional state, the separable mixture model (SMM) [Hofmann and Puzicha 1998] is suitable for modeling the co-occurrence data trained to evaluate the conditional probability of the emotional states given by the SLs and ATTs. In the test phase, the test sentence is first transferred into the SL and ATT sequences by applying the same universal lexical ontology. The SMM is then adopted to estimate the maximal probability of the emotional state.

In the training and test phases, the universal lexical ontology is important for providing prior knowledge of the input sentences. In the ontology, each concept is represented by a set of attributes (ATTs) which are used to describe the hierarchical relationship between the concepts. The ontology used in our system was developed by Chen et al. [2003] based on the Chinese knowledge base, HowNet [Dong and Dong 1988], to represent the lexical knowledge. To expand the depth and width of the concept relationships, WordNet [Fellbaum 1998] was integrated into the ontology using a bilingual dictionary. Every concept in the universal lexical ontology will finally relate to several attributes (ATTs) via the hierarchical links.

3. EMOTION GENERATION RULE

Further textual data analysis of emotion recognition shows that not only keywords convey the emotion but the general terms do also. For example, a speaker may say "I finally finished that annoying job" instead of "*I am so glad that* I finally finished that annoying job." The two sentences express the same emotional state, but the critical emotional keyword "*glad*" is not spoken in the first sentence. To solve this problem, the mechanism for generating emotional states must first be investigated.

The conditions for generating emotions are summarized according to the previous research on emotion psychology [Lazarus and Lazarus 1996]. These conditions are

Table I. Some Examples of EGRs

EGRs	Domain-independent component	Domain-dependent component
One may be HAPPY *if he reaches his goal*	[REACH]/[OBTAIN]/[CLOSE TO]	Something that can be a goal
One may be HAPPY *if he have someone's support*	[REACH]/[OBTAIN]/[CLOSE TO]+ [SHOW GOOD EMOTION] / [POSITIVE EVALUSTION]	Something that can be a kind of support
One may be HAPPY *if he loses something harmful*	[LOSE]/[FAR AWAY]	Something harmful

assumed to be generic even though different individual backgrounds and environments are involved. For example, we may feel happy when we have a good reputation, reach a high position in our job, or obtain any kind of benefit. Then "*One may be HAPPY if he obtains something beneficial*" will be one of the generation rules for a happy emotion. These kinds of conditions or environmental situations are based on emotion psychology and are manually derived as emotion generation rules (EGRs) in this article.

Although EGRs are able to describe situations producing specific emotional states, there still exist some ambiguities inside EGRs. Take the previous EGR as an example; it is clear that "*One may be HAPPY if he obtains something beneficial*," but the emotional state of "someone lost something beneficial" may be ANGRY or SAD, according to his/her personal characteristics. Accordingly, to eliminate the ambiguities in EGRs, we deduce EGRs with only two opposite emotional states: HAPPY and UNHAPPY.

Each EGR, as mentioned above, consists of domain-independent and domain-dependent components. Generally speaking, domain-independent components are "verbs" in EGRs, while domain-dependent components are "nouns" in EGRs. As in the above example, "*obtains*" and "*something beneficial*" are respectively the domain-independent and domain-dependent components of the EGR "*One may be HAPPY if he obtains something beneficial*." Table I shows some examples of EGRs and the corresponding emotional state, domain-independent, and domain-dependent components.

4. SEMANTIC LABEL DEFINITION AND EXTRACTION

In the above EGR, the definition and extraction of SLs are critical to the entire process. The semantic label, as the literal meaning, is defined as a word or a phrase that indicates some specific semantic information. Due to their domain-independent characteristics, it is

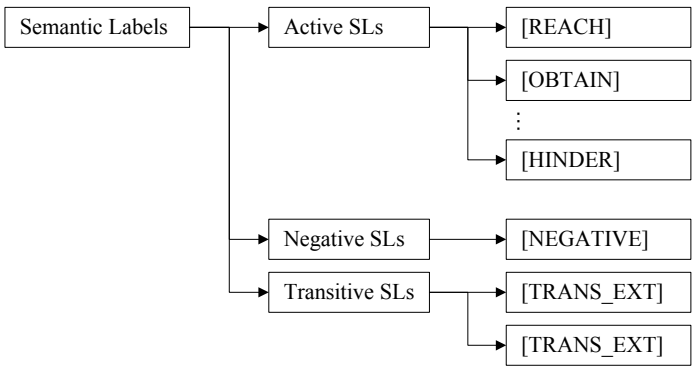

Fig. 2. The SL tree structure.

reasonable to define them with the universal lexical ontology. This section introduces the SL definition and extraction method. The resulting SL tree structure is illustrated in Figure 2.

4.1 Semantic Label Definition

In Figure 2, three kinds of SLs are defined: active SLs, negative SLs, and transitive SLs. With the help of hypernyms defined in the universal lexical ontology, we first define the active SLs which form the main EGR intention, such as [REACH], [OBTAIN], or [LOSE]. Table II shows 15 currently defined active SLs. Because most of the concepts that have the same intention share the same hypernyms in the ontology, we can define them by simply defining their hypernyms. For example, the concept "receive" is the hypernym of concepts "get," "acquire," "obtain," and "earn." With the definition "*concepts that have a hypernym* "receive" *can be the SL* [OBTAIN]," we can convert the four words "get," "acquire," "obtain," and "earn" in the input sentence to the SL [OBTAIN]. Some definition of active SLs can also be hierarchical; that is, the definition of an SL may be the hypernym of a concept's hypernym. The active SL definition job can only be made manually, and was actually done by three researchers in about three months. As the result, 147 hypernyms in total were selected from 803 hypernyms for the definitions of 15 active SLs.

It is simpler for the definition of two other kinds of SLs than for active SLs. The negative SLs indicate the words with negative semantic meanings. Only one SL is contained in this kind of SL: [NEGATIVE]. The transitive SLs indicate the transitive semantic meaning and contain two SLs: [TRANS_EXT] and [TRANS_EXT]. The transitive-extractive SL [TRANS_EXT] imply that the ensuing sentence is more important than the preceding sentence, so the SLs before [TRANS_EXT] will be ignored. Conversely, the SLs following the transitive-subtractive SL [TRANS_SUB] will be ignored. Because the concepts belonging to these two SLs are enumerable, it is not necessary for these two SLs to define hypernyms hierarchically in the ontology. For the negative SLs, 47 words are directly defined as [NEGATIVE] SLs, such as "cannot," "unnecessary," "no," "never," and so on. In the transitive SLs, 18 and 11 words can be directly defined as [TRANS_EXT] and [TRANS_SUB], respectively. The corresponding words of SL [TRANS_EXT] include "finally," " but," "fortunately," and so on. And the corresponding words of SL [TRANS_SUB] include "or," "even though," "although,"and so on..

4.2 Automatic SL Extraction

The previous three kinds of SLs can be automatically extracted after they are defined. The reason for automatic SLs extraction is to improve system portability. Since we assumed that the emotion reaction to the same sentences depends on personal status, the emotion classification model must be retrained for a new personal status. The emotional

Table II. The 15 Defined Active SLs

EGRs	Domain-independent component	Domain-dependent component
One may be HAPPY *if he reaches his goal*	[REACH]/[OBTAIN]/[CLOSE TO]	Something that can be a goal
One may be HAPPY *if he have someone's support*	[REACH]/[OBTAIN]/[CLOSE TO]+ [SHOW GOOD EMOTION] / [POSITIVE EVALUSTION]	Something that can be a kind of support
One may be HAPPY *if he loses something harmful*	[LOSE]/[FAR AWAY]	Something harmful

corpus must be re-collected and re-annotated. Due to the subjective presumption, annotating all of the appropriate SLs for the corpus is time-consuming and always imprecise.

As mentioned above, the SL definitions are different for different SLs. When a textual sentence is input, the negative SLs and transitive SLs are first converted by simply comparing the predefined words. Then, the hierarchical hypernym structure of the remaining words in the sentence is checked to find appropriate active SLs. When a word can match more than one SL, all of them will be retained for the following training process.

5. GENERATING RULES FOR EMOTION ASSOCIATION

So far we have defined the SLs that represent the domain-independent components of EGRs which can be extracted automatically. To determine the domain-dependent components in EGRs, the emotion association rule (EAR) training process is described in this section. After SL extraction, those words that are not converted into SLs are first converted into their corresponding ATTributes (ATTs), which are defined in the lexical ontology for concept representation. The rule-generation algorithm is then applied to mine the association between the SL and ATT sets with high co-occurrence. The uncorrelated ATTs are discarded. The *a priori* algorithm is adopted as the rule-generation algorithm, which is generally used to mine association rules from raw data. The following sections introduce the *a priori* algorithm and the EARs' generation process.

5.1 The *A Priori* Algorithm

The *a priori* algorithm was proposed in 1994 by Agrawal to mine the association rules from transaction data [Han and Kamber 2001]. The transaction data is a list of item-sets. Each item-set contains an arbitrary number of items selected from a finite set of items. Two important parameters are introduced first: support and confidence. Calculating these two parameters is described in Eqs.(1) and (2).

$$Sup(A,B) = \frac{T(A,B)}{T_{total}}\ , \tag{1}$$

$$Conf(A \rightarrow B) = \frac{T(A,B)}{T(A)}\ . \tag{2}$$

Where A and B are two item-sets, $Sup(A,B)$ and $Conf(A \rightarrow B)$ are the support and confidence functions, respectively. $T(A)$ is the amount of A in all transactions and $T(A,B)$ is the amount of co-occurrences of A and B in all transactions. The value T_{total} is the total number of transactions. The union of A and B, which is also an item-set, is called "large item-set" if both $Sup(A,B)$ and $Conf(A \rightarrow B)$ are larger than the predefined thresholds.

Basically, the *a priori* algorithm procedure involves finding all large item-sets with variant item numbers. The redundant item-sets are then pruned. Each one of the large item-sets stands for one association rule and the remaining large item-sets are the final resulting association rules. The pruning procedure is described in the following example. Assume that there are two large item-sets $\{\{i_1, i_2\}, i_3\}$ and $\{i_1, i_3\}$. The corresponding rules are: "*If the items i_1 and i_2 appear, then the item i_3 will appear*" and "*If the item i_1 appears, then the item i_3 will appear*," respectively. It is obvious that the first rule can be pruned because it is a subset of the second rule.

The pruning procedure can accelerate the *a priori* algorithm as well as improve the quality of the association rules. However, in our case, an additional problem occurs when a word is converted into a set of ATTs: the ATT sets will appear in all sentences, and the

value of *Sup* and *Conf* of these ATT sets are always larger than the predefined threshold for "large item-set." That means the ATT set for a word will always be an EAR. Because they are extracted from one word, to extract these kinds of rules is meaningless. To solve this problem, a new restriction is added to the original pruning procedure: all large item-sets containing items that can be converted from the same word will also be pruned. This new restriction is tested at each step to ensure that there are no redundant large item-sets.

5.2 Emotion Association Rule Generator

The emotion association Rrules (EARs) used to represent the association between SLs and ATTs for a specific emotional state are derived from the training text corpus via the *a priori* algorithm. The process for deriving the EARs follows:

Step 1. Sentence preprocessing:
preprocess each training sentence such as word segmentation, stop-word deletion, and so on;

Step 2. SL extraction:
compare the hypernyms of each word in the training sentence with the predefined SLs. If the hypernyms of a word match the SLs, mark the word as the corresponding SL.

Step 3. SL and ATT set generation:
convert the remaining words into their corresponding ATTs, except for the words marked as SLs. Repeat steps 1 to 3 until all training sentences are converted into SLs and ATTs;

Step 4. EARs derivation:
For each emotional state, apply the *a priori* algorithm to derive EARs from all the SLs and ATTs obtained from the training sentences tagged as the corresponding emotional state.

The following example describes the steps in the extraction process:

Input sentence
I got the electricity bill; it is actually less than 100 dollars!

Step 1. Sentence preprocessing
➔ *got electricity bill; actually less than 100 dollars!*

Step 2. SLs extraction
get (*hypernyms* = *receive* or *fulfil*): [OBTAIN] or [REACH]
actually: [TRANS_EXT]
less than: [NEGATIVE]
➔ [OBTAIN] *electricity bill;* [NEGATIVE] *100 dollars!*
➔ [REACH] *electricity bill;* [NEGATIVE] *100 dollars!*

Step 3. Generation of SL and ATT sets
power: DEF= expenditure, #electricity
bill: DEF= bill, #wealth
100 dollars: [NUM]
➔ [NEGATIVE], [NUM], (*expenditure, #electricity*), (*bill, #wealth*)

Step 4. EARs derivation (some examples)
(*repeat step1 to step3 until all sentence in corpus are converted into sets of SLs and ATTs*)
Happy: [SHOW GOOD EMOTION] + [CLOSE TO] + human → HAPPY
Happy: [CLOSE TO] + human + request → HAPPY
Unhappy: [NEGATIVE] + [NUM] + bill → UNHAPPY

In this example the input sentence is preprocessed first. In step 2, the appropriate SLs are extracted, including active SLs, negative SLs, and transaction SLs. When deciding the appropriate active SLs for the input sentence, we first determine the POS of the phrases. For each possible word, the corresponding ATT is compared with the SLs. If a word matches more than one SL, all the SLs will be kept for further processing. In this example, "*receive*" and "*fulfil*," which are two hypernyms of the verb "*got*," are the definitions of the SLs [OBTAIN] and [REACH], respectively. Therefore, two sentences are used in the second step of the process. Except for the active SLs, two phrases "*actually*" and "*less than*" are important keywords that match the transitive SLs and negative SLs, respectively. Due to the assumption that the word "*actually*" is used to emphasize the sentence after it, the two active SLs [OBTAIN] and [REACH] are then filtered out. In this step all of the phrases are converted into their corresponding SLs. In step 3, the SL and ATT sets are constructed by converting the remaining words into their ATTs. Furthermore, the phrase "*100 dollars*" is tagged as [NUM] instead of the original ATTs. The final result is shown in step 3, and is used for further processing in step 4. The EARs are then deduced using the *a priori* algorithm.

6. THE SEPARABLE MIXTURE MODEL

The most important process for the test phase is to determine the emotional state of the input sentence, that is, to identify the similarity between the input sentence and the EARs of a specific emotional state represented by SLs and ATTs. Because each input sentence may match more than one EAR for different emotional states, the separable mixture model (SMM) [Hofmann and Puzicha 1998] is appropriate for the classification process and for emotion recognition. Given the input sentence s, the corresponding emotional state is decided by maximizing the conditional probability of each emotional state e_k.

$$e^* = \arg\max_{e_k} P(e_k|s). \tag{3}$$

According to the assumption that the emotional state of an input sentence can be determined using the EAR with respect to the input sentence, this conditional probability can be expanded as

$$P(e_k|s) = \sum_{r=1}^{R} P(e_k|EAR_r)P(EAR_r|s).$$

The variable EAR_r in this equation indicates the r^{th} EAR. From the description in previous sections, the probability $P(EAR_r|s)$ is defined as 1 if EAR_r can be derived from the input sentence s. Otherwise, $P(EAR_r|s)$ is set to 0.

To estimate the distribution $P(e_k|EAR_r)$, we assume that $EAR_r = \{SL_1, \ldots, SL_{m(r)}, ATT_1, \ldots, ATT_{n(r)}\}$ is the r^{th} EAR where $m(r)$ and $n(r)$, respectively, are the number of SLs and ATTs in this EAR. The conditional probability can be further expanded according to the Bayes' rule:

$$P(e_k|EAR_r) \propto P(EAR_r|e_k) = P(SL_1,\ldots,SL_{m(r)}, ATT_1,\ldots,ATT_{n(r)}|EAR_r).$$

Since every EAR is collected using some fixed threshold, the conditional probability of an emotional state given by an obvious EAR tends to be smaller than the actual probability. If we treat $P(EAR_r|e_k)$ as a distribution, the emotional state of the input

sentence will be the summation of all distributions. To estimate this distribution, the separable mixture model is adopted.

The original purpose of SMM is to model the appearance of co-occurrence data (x, y), where x and y are selected from two categories $\boldsymbol{X}$ and $\boldsymbol{Y}$. In our application, since each EAR is made up of several SLs and ATTs, the appearance of EAR_r means the simultaneous appearance of all of the SLs and ATTs: $\{SL_1, \ldots, SL_{m(r)}, ATT_1, \ldots, ATT_{n(r)}\}$. By assuming that the appearance of the SLs and ATTs is conditionally independent, SMM can be adopted to fit our problem.

The conditional probability of EAR_r is expanded as

$$\begin{aligned} P\left(EAR_r \middle| e_k\right) &= P\left(SL_1,\ldots,SL_{m(r)},ATT_1,\ldots,ATT_{n(r)} \middle| e_k\right) \\ &= \alpha_k \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} P\left(SL_i \middle| e_k\right) P\left(ATT_j \middle| e_k\right) \quad , \\ &\equiv \alpha_k \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|k} q_{j|k} \end{aligned}$$

where α_k is the probability of selecting the emotional state e_k; and $p_{i|k}$ and $q_{j|k}$ are the conditional probabilities of selecting SL_i and ATT_j based on the emotional state e_k,. Therefore, the distribution of EAR_r is the summation of $P(EAR_r|e_k)$.

$$P\left(EAR_r\right) = \sum_{k=1}^{K} \left(\alpha_k \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|k} q_{j|k} \right).$$

The parameters of this distribution can be estimated using the EM algorithm [McLachlan and Krishnan 1997]. During the E-step, the expectation of $P(EAR_r)$ is

$$L = \sum_{r=1}^{R} \beta_r \log \left(\sum_{c=1}^{K} \left(\alpha_c \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|c} q_{j|c} \right) \right), \tag{4}$$

where β_r is the number of EAR_r in the corpus. A new variable R_{rk} is then introduced as a Boolean variable; $R_{rk} \in \{0,1\}$ indicates whether EAR_r belongs to the specific emotional state e_k. By involving this parameter, the expectation in Eq. (4) is simplified to

$$L' = \sum_{r=1}^{R} \sum_{k=1}^{K} R_{rk} \log \left(\alpha_k \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|k} q_{j|k} \right) = \sum_{r=1}^{R} \sum_{k=1}^{K} R_{rk} \left(\log \alpha_k + \sum_{i=1}^{m(r)} \log p_{i|k} + \sum_{j=1}^{n(r)} \log q_{j|k} \right).$$

With a similar assumption of SMM in Hofmann and Puzicha [1998], the representation and normalization constraint of the introduced parameter R_{rc} is shown in Eqs. (5) and (6), respectively:

$$R_{rk}^{(t+1)} = \frac{\alpha_k^{(t)} \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|k}^{(t)} q_{j|k}^{(t)}}{\sum_{v=1}^{K} \left(\alpha_v^{(t)} \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v}^{(t)} q_{j|v}^{(t)} \right)}, \tag{5}$$

$$G = \left(\sum_{r=1}^{R} \sum_{k=1}^{K} R_{rk} \right) - R = 0, \tag{6}$$

Due to this constraint, the Lagrange multiplier is used herein to obtain the maximum partial differentiation. In the following formulas, we take the partial differentiation of α_x as an example.

$$\frac{\partial}{\partial \alpha_x} L + \lambda \frac{\partial}{\partial \alpha_x} G = 0$$
$$\Rightarrow \frac{\partial}{\partial \alpha_x} \sum_{r=1}^{R} \sum_{k=1}^{K} \left[R_{rk} \left(\log \alpha_k + \sum_{i=1}^{m(r)} \log p_{i|k} + \sum_{j=1}^{n(r)} \log q_{j|k} \right) \right] + \lambda \frac{\partial}{\partial \alpha_x} \sum_{r=1}^{R} \sum_{k=1}^{K} R_{rk} = 0 \quad , \qquad (7)$$

In this equation, the term $\frac{\partial}{\partial \alpha_x} R_{rk}$ is deduced for both cases: $k = x$ and $k \neq x$.

$$\frac{\partial}{\partial \alpha_x} R_{rx} = \frac{\partial}{\partial \alpha_x} \frac{\alpha_x \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x}}{\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right)}$$
$$= \frac{\left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right) \left(\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right) - \left(\alpha_x \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right) \left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right)}{\left[\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right]^2}$$
$$= \frac{\left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right) \left(\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right)}{\left[\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right]^2} - \frac{\left(\alpha_x \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right) \left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right)}{\left[\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right]^2}$$
$$= \frac{R_{rx}}{\alpha_x} - \frac{R_{rx}^2}{\alpha_x}$$

$$\frac{\partial}{\partial \alpha_x} R_{ry} = \frac{\partial}{\partial \alpha_x} \frac{\alpha_y \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|y} q_{j|y}}{\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right)} = \frac{-\left(\alpha_y \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|y} q_{j|y} \right) \left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right)}{\left[\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right) \right]^2}$$
$$= \left[\frac{-\left(\alpha_y \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|y} q_{j|y} \right)}{\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right)} \right] \left[\frac{\left(\prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|x} q_{j|x} \right)}{\sum_{v=1}^{K} \left(\alpha_v \prod_{i=1}^{m(r)} \prod_{j=1}^{n(r)} p_{i|v} q_{j|v} \right)} \right]$$
$$= \frac{-R_{rx} R_{ry}}{\alpha_x}$$

The previous two equations represent the result for the $k = x$ and $k \neq x$ cases. The result is then applied to Eq. (7). After expanding all the differentiation terms, the first partial differentiation result is obtained.

$$\sum_{r=1}^{R}\left(\log\alpha_x+1+\sum_{i=1}^{m(r)}\log p_{i|x}+\sum_{j=1}^{n(r)}\log q_{j|x}+\lambda\right)-\sum_{r=1}^{R}\sum_{k=1}^{K}\left(R_{rk}\left(\log\alpha_k+\sum_{i=1}^{m(r)}\log p_{i|k}+\sum_{j=1}^{n(r)}\log q_{j|k}+\lambda\right)\right)=0$$

The other two partial differentiation results for the expectation equation can also be obtained in the same way. The iterative equations for the parameters are resolved using the following three simultaneous linear equations.

$$\alpha_k^{(t)} = \frac{1}{R}\sum_{r=1}^{R}R_{rk}^{(t)}$$

$$p_{x|k}^{(t)} = \frac{1}{R\alpha_k^{(t)}}\times\frac{\sum_{r=1}^{R}R_{rk}^{(t)}}{\sum_{i(r)\neq x}p_{x|k}^{(t-1)}} \quad , \quad 1\le x\le m(r)$$

$$q_{y|k}^{(t)} = \frac{1}{R\alpha_k^{(t)}}\times\frac{\sum_{r=1}^{R}R_{rk}^{(t)}}{\sum_{j(r)\neq y}q_{y|k}^{(t-1)}} \quad , \quad 1\le y\le n(r)$$

The EAR distribution can be iteratively estimated and the appropriate emotional state for a given input sentence e^* can be determined using Eq. (3). In practice, an additional value δ is experimentally defined, and all sentences with the SMM likelihood less than δ will be rejected.

7. EXPERIMENTAL RESULTS

Because ATTs are defined as domain-dependent, we consider a limited application domain for performance evaluation. A dialog system focusing on the students' daily expressions was constructed, and only three emotional states, HAPPY, UNHAPPY, and NEUTRAL, were considered. To evaluate the proposed approach, four experiments were conducted. Experiment 1 was to determine the best parameter combination for the entire system. In experiment 2 the parameter set was to test the accuracy of emotion recognition. In experiment 3 we ported the system into another domain and evaluated recognition accuracy. Comparisons to other approaches were investigated in experiment 4.

Two dialog corpora were collected for the following experiments. Corpus A consisted of 26 volunteer college students. To counterbalance day-to-day variations, we continually collected data for one month. The data collection system was basically a dialog system that could guide speakers into talking about their daily lives by asking questions and recording the answers. The initial questions were selected randomly from 13 predefined question sets; each question focused on one topic in the student's daily life. According to the previous training for emotional representation, the students were further divided into two groups. In Group 1 with 6 students, each student recorded 40 dialogs for each emotion. For each of the remaining 20 students, 10 dialogs were recorded for each emotion. Corpus B contains the dialogs from a broadcast drama. In both corpora, we annotated happy, unhappy, and neutral emotional states for each sentence. All sentences not belonging to these three emotional states were deleted from the corpus. The dialogs of different performers were extracted and the emotions of the performers were manually tagged with one of the above three emotional states. Only the protagonists' dialogs were preserved in the corpus. The sentences with three words or less were also discarded. The content of these two corpora are shown in Tables III and IV.

Table III. Corpus A Content

	Corpus A-1	Corpus A-2
Number of speakers	6	20
Number of one-sentence dialogs	720	300
Number of multi-sentence dialogs	720	300
Number of happy dialogs	480	200
Number of unhappy dialogs	480	200
Number of neutral dialogs	480	200
Total number of dialogs	1440	600

Table IV. Corpus B Content

Number of speakers	7
Number of happy dialogs	283
Number of unhappy dialogs	594
Number of neutral dialogs	1225
Total number of dialogs	2102

7.1 Determining Parameter Combinations

In the first experiment, corpus A-1 was used to estimate the best parameter combination. From the description in previous sections, the emotion recognition process includes SL extraction, EAR generation, and SMM construction. The SL extraction procedure is based on predefined knowledge and has no parameter for tuning. The sentences that cannot complete the transformation are marked as neutral emotion. Table V shows the results of SL and ATT set transformations and the recall rates of neutral emotions. In the EAR generation procedure, the parameters include *Sup* (support) and *Conf* (confidence). In the SMM construction procedure, we conducted the experiment using an additional parameter, δ. The parameter δ is a threshold for SMM in which all sentences with SMM likelihood less than δ are rejected. These three parameters are mutually related and directly affect the performance of the recognition system. Hence we cannot independently modify any one of them.

The four contours of the results are shown in Figure 3 with different mark types: diamond, square, triangle, and cross corresponding to different *Sup* values: 0.5, 1, 1.5, and 2, respectively. From these figures, we find that the best results are obtained when $\delta = 0.1$, Sup = 1, and Conf = 5.

Table V. SL and ATT Sets and Recognition Accuracy Results for Neutral Emotions in Corpus A-1

	Happy	Unhappy	Neutral
Number of original dialogs	480	480	480
Number of SL and ATT sets	404	422	84
Number of remaining dialogs	76	58	396
Precision rate of neutral emotion	$\frac{396}{76+58+396} = 0.7472$		
Recall rate of neutral emotion	$\frac{396}{480} = 0.825$		

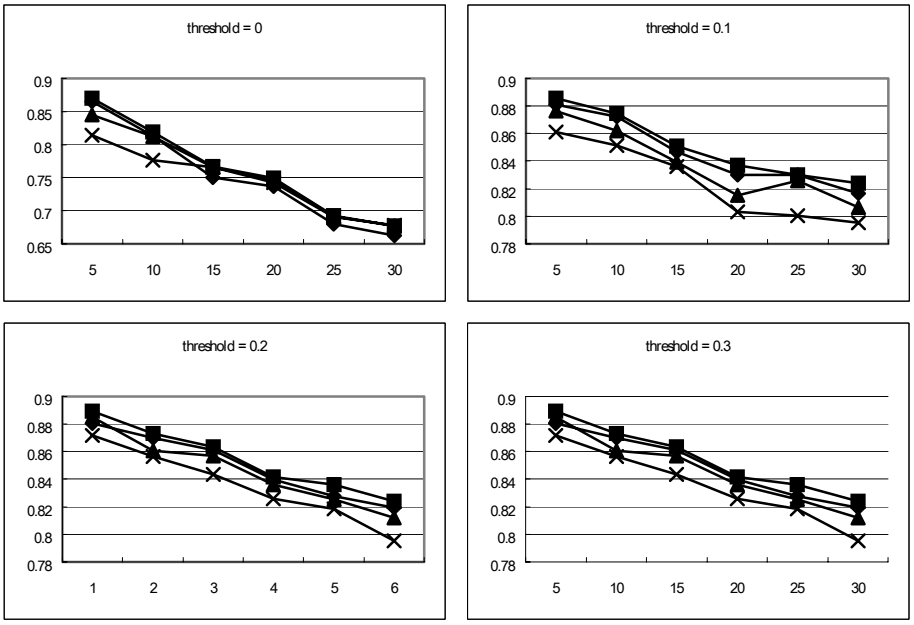

Fig. 3. Recognition accuracy with different parameter sets. The lines represent different values of *Sup*. X- and Y-coordinates represent the values of *Conf* and accuracy, respectively.

7.2 The Accuracy of Emotion Recognition

Following the previous parameter combination and recognition process, corpus A is used to evaluate recognition accuracy. Corpus A-1 is used for inside testing and Corpus A-2 for outside testing. Table VI shows the inside test results, from which we can calculate the precision, recall, and rejection rates for the three groups of emotions in Table VII.

From the results we find that the precision rate can achieve 83.94% under a recall rate of 75.90%. The rejection rate is 5% on average, except for neutral emotions. The reason for the high rejection rate for the neutral emotions is that we tend to extract as many SLs as possible. Many sentences with neutral emotion can be converted to sets of SLs and ATTs. However, in the next step, with no SLs necessary, many SL and ATT sets were rejected using the threshold.

Table VI. Inside Test Experimental Results

	Happy	Unhappy	Neutral
	480	480	480
process 1	422	404	84
process 2	408	379	42
Happy	364	46	26
Unhappy	44	333	16

Number of neutral dialogs		
Happy	Unhappy	Neutral
58	76	396
Number of rejections		
Happy	Unhappy	Neutral
14	25	42

Table VII. Precision, Recall, and Rejection Rates in Table VI

	Happy	Unhappy	Neutral	Average
Precision rate (%)	88.78%	88.33%	74.72%	83.94%
Recall rate (%)	75.83%	69.38%	82.50%	75.90%
Rejection rate (%)	3.32%	6.19%	33.33%	14.28%

Table VIII. Outside Test Experimental Results

	Happy	Unhappy	Neutral
	200	200	200
process 1	173	166	56
process 2	160	141	21
Happy	132	23	12
Unhappy	28	118	9

Number of neutral dialogs		
Happy	Unhappy	Neutral
27	34	144
Number of rejections		
Happy	Unhappy	Neutral
13	25	35

Table IX. Precision, Recall, and Rejection Rates in Table VIII

	Happy	Unhappy	Neutral	Average
Precision rate (%)	79.04%	76.13%	70.24%	75.14%
Recall rate (%)	66.00%	59.00%	72.00%	65.67%
Rejection rate (%)	7.51%	15.06%	62.50%	28.36%

Table VIII shows the outside test results; Table IX shows the calculation process. Even with outside testing, the precision rate still achieved 75.14% under a recall rate of 65.67%, which shows that the proposed approach is satisfactory for a specific domain.

7.3 Portability Evaluation

To evaluate the portability of the system, corpus B was used in this experiment. Because corpus B was collected from a broadcast drama performed by professional actors, the emotion reaction in this corpus was more realistic than that of corpus A. Before porting the original system into this new corpus, the only work was annotating the emotion reaction in the sentences. Our approach assumed that only the primary emotional states, such as happy or unhappy, were necessary for annotation.

Table X. Portability Test Experimental Results

	Happy	Unhappy	Neutral
	283	594	1225
process 1	173	214	242
process 2	104	163	121
Happy	92	28	72
Unhappy	12	135	49

Number of neutral dialogs		
Happy	Unhappy	Neutral
110	380	983
Number of rejections		
Happy	Unhappy	Neutral
69	51	121

Tables X and XI show the results of the porting process. Only half of the sentences in corpus B could be converted into the SL and ATT sets. The final precision rate was 61.18% under a recall rate of 45.16%. Because the broadcast drama was not a domain-specific corpus, the proposed approach could not be applied directly very well. However, in another simple experiment, when we focused only on one actor, the precision rate reached 70%.

Table XI. Precision, Recall, and Rejection Rates in Table X

	Happy	Unhappy	Neutral	Average
Precision rate (%)	47.92%	68.88%	66.73%	61.18%
Recall rate (%)	32.51%	22.73%	80.24%	45.16%
Rejection rate (%)	39.88%	23.83%	50.00%	37.91%

7.4 Comparison to Other Systems

Several emotion recognition systems were introduced in the previous sections. In this experiment, two different approaches, semantic network and emotional keywords, are compared. In approach 1, the details for constructing the semantic network are described in Woods [1970], where the emotional state of an input dialog is identified by considering its syntactic structure and the relationships between lexicons. In approach 2, the emotional keywords are selected manually from the universal lexical ontology. The relationships between the keywords and emotional groups are then calculated statistically.

Table XII. Comparing Classification Accuracy

	Precision rate (%)		
	Corpus A-1	Corpus A-2	Corpus B
Approach 1 (Semantic Net)	76.42%	70.01%	55.39%
Approach 2 (Emotional keyword)	82.15%	73.44%	32.97%
The proposed approach	80.98%	75.14%	61.18%

Using both corpora A and B, Table XII shows the average accuracies for the two systems. In the inside test, using a small and specific domain corpus, the keyword-based system is better than the other two. However, when ported into other domains, the other two systems achieved better performance. Generally speaking, the result from the proposed approach is better than the other two systems in a normal situation, even though there were several sentences in a dialog. The result shows the advantage of applying psychology and the separable mixture model.

8. CONCLUSION

An automatic emotion recognition approach from text was presented in this article. By applying a universal lexical ontology and research on the psychology of emotion, we defined the EGRs and SLs to automatically extract semantic information from input sentences; the SLs and ATTs represent implicit semantic information in EGRs. The sentences were converted into the SL and ATT sequences, and the *a priori* algorithm was applied to determine the EARs. To identify the relationships between new sentences and EARs, a separable mixture model was adopted to estimate the probability of the emotional state given in the input sentence. The experimental results show that the proposed approach achieves better performance at emotion recognition in a dialog than the other systems we measured. With the combination of semantic information and psychology, the proposed approach identifies the correct emotional state in a specific domain, and works well in other domains as well. Due to the ambiguities in EGRs, we adopted only happy, unhappy, and neutral emotional states for evaluation. Our future research using this approach will examine the emotion history effect. The expansion of the universal lexical ontology will be a great help in extracting semantic information.

REFERENCES

AMIR, N., ZIV, S., AND COHEN, R. 2003. Characteristics of authentic anger in Hebrew speech. In *Proceedings of the 8th European Conference on Speech Communication and Technology* (Geneva). 713-716.

BOUCOUVALAS, A.C. 2002. Real time text-to-emotion engine for expressive internet communications. In *Emerging Communication: Studies on New Technologies and Practices in Communication*. G. Riva et al. eds. IOS Press. 305-318.

BOUCOUVALAS, A.C., AND ZHE, X. 2002. Text-to-emotion engine for real time internet communication. In *Proceedings of the International Symposium on CSNDSP 2002* (Staffordshire Univ., July 15-17). 164-168.

CHAN, S.W.K. AND FRANKLIN, J. 1998. Symbolic connectionism in natural language disambiguation. *IEEE Tran. Neural Network. 9*, 739-755.

CHEN, M.J., YEH, J.F., AND WU, C.H. 2003. Ontology-based dialog management for service integration. In *Proceedings of ROCLING XV* (Hsinchu, Taiwan). 257-277.

CHUANG, Z.J. AND WU, C.H. 2002. Emotion recognition from textual input using an emotional semantic network. In *Proceedings of the International Symposium on Chinese Spoken Language Processing* (Denver, CO). 177-180.

COHN, J.F. AND KATZ, G.S. 1998. Bimodal expression of emotion by face and voice. In *Proceedings of the Sixth ACM International Multimedia Conference on Face/Gesture Recognition and Their Applications* (Bristol, UK). ACM, New York. 41-44.

COWIE, R., DOUGLAS-COWIE, E., TSAPATSOULIS, N., VOTSIS, G., KOLLIAS, S., FELLENZ, W., AND TAYLOR, J.G. 2001. Emotion recognition in human-computer interaction. *IEEE Signal Process. Mag. 18,* 1*,* 32-80.

DE SILVA, L.C. AND NG, P.C. 2000. Bimodal emotion recognition. In *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000* (March). 332-335.

DEVILLERS, L., VASILESCU, I., AND LAMEL, L. 2002. Annotation and detection of emotion in a task-oriented human-human dialog corpus. In *Proceedings of the ISLE Workshop on Dialogue Tagging for Multi-Modal Human-Compute Interaction* (Dec. 15-17).

DEVILLERS, L., LUNIEL, L., AND VASILESCU, I. 2003. Emotion detection in task-oriented spoken dialogues. In *Proceedings of the International Conference on Multimedia and Expo* (Baltimore, MD, July 6-9). 549-552.

DIJKSTRA, K., ZWAAN, R.A., GRAESSER, A.C., AND MAGLIANO, J.P. 1994. Character and reader emotions in literary texts. *Poetics 23*, 139-157.

DONG, Z., AND DONG, Q. 1988. HowNet. http://www.keenage.com/zhiwang/e_zhiwang.html.

FELLBAUM, C. 1998. *WordNet: An Electronic Lexical Database*. MIT Press. Cambridge, MA.

FORBES-RILEY, K. AND LITMAN, D.J. 2004. Predicting emotion in spoken dialogue from multiple knowledge sources. In *Proceedings of Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics* (HLT/NAACL, Boston, MA, May 2-7). 201-208.

HAN, J. AND KAMBER, M. 2001. *Data Mining: Concepts and Techniques*. Morgan Kaufmamn.

HOFMANN, T. AND PUZICHA, J. 1998. Statistical models for co-occurrence data. AI Memo 1625; CBCL Memo159, Feb.

KWON, O.W., CHAN, K., HAO, J., AND LEE, T.W. 2003. Emotion recognition by speech signals. In *Proceedings of the 8th European Conference on Speech Communication and Technology* (Geneva, Sept.). 125-128.

LAZARUS, R.S. AND LAZARUS, B.N. 1996. *Passion and Reason: Making Sense of Our Emotions*. Oxford University Press, New York.

LEE, C.M., NARAYANAN, S.S., AND PIERACCINI, R. 2002. Combining acoustic and language information for emotion recognition. In *Proceedings of the 7th International Conference on Spoken Language Processing* (Denver, CO). 873-876.

LITMAN, D. AND FORBES, K. 2003. Recognizing emotions from student speech in tutoring dialogues. In *Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop* (Virgin Islands, Dec.). 39-52.

MCLACHLAN, G.J. AND KRISHNAN, T. 1997. *The EM Algorithm and Extensions*. Wiley, New York.

PICARD, R.W. 1997. *Affective Computing*. MIT Press. Cambridge, MA.

PICARD, R.W., VYZAS, E., AND HEALEY, J. 2001. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Trans. Pattern Anal. Mach. Intell. 10,* 1175-1191.

RAHURKAR, M.A. AND HANSEN, J.H.L. 2003. Frequency distribution based weighted sub-band approach for classification of emotional/stressful content in speech. In *Proceedings of the 8th European Conference on Speech Communication and Technology* (Geneva, Sept. 1-4). 721-724.

SCHULLER, B., RIGOLL, G., AND LANG, M. 2004. Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture. In *Proceedings of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Montreal, May 17-21). 17-21.

SUBASIC, P. AND HUETTNER, A. 2001. Affect analysis of text using fussy semantic typing. *IEEE Trans. Fuzzy Syst. 9*, 483-496.

TAO, J. AND TAN, T. 2004. Emotional Chinese talking head system. In *Proceedings of the 6th International Conference on Multimodal Interface* (Oct. 13-15). 273-280.

WOODS, W.A. 1970. Transition network grammars for natural language analysis. *Commun. ACM 13*. 591-602.

YANARU, T. 1995. An emotion processing system based on fuzzy inference and subjective observations. In *Proceedings of the 2nd New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems* (Dunedin, N Z., Nov. 20-23). IEEE Computer Society Press, New York, 15-20.

YU, F., CHANG, E., XU, Y.Q., AND SHUM, H.Y. 2001. Emotion detection from speech to enrich multimedia content. In *Proceedings of the Second IEEE Pacific-Rim Conference on Multimedia* (Beijing, Oct.24-26).

Received March 2004; revised March 2005; accepted February 2006.