An Automated Crisis Online Dispatcher

Siska Fitrianie
Man-Machine Interaction Group,
Delft University of Technology
s.fitrianie@ewi.tudelft.nl

Leon J.M. Rothkrantz
Man-Machine Interaction Group,
Delft University of Technology
l.j.m.rothkrantz@ewi.tudelft.nl

ABSTRACT

An experimental automated dialogue system that plays the role of a crisis hotline dispatcher is currently developed. Besides controlling the communication flow, this system is able to retrieve information about crisis situations from user's input. It offers a natural user interaction by the ability to perceive and respond to human emotions. The system has an emotion recognizer that is able to recognize the emotional loading from user's linguistic content. The recognizer uses a database that contains selected keywords on a 2D "arousal" and "valence" scale. The output of the system provides not only the information about the user's emotional state but also an indication of the urgency of his/her information regarding to crisis. The dialogue system is able to start a user friendly dialogue, taking care of the content, context and emotional loading of user's utterances.

Keywords

Crisis management, human-computer dialogue system, natural language processing, text-based emotion recognition.

INTRODUCTION

Automatic dialogue systems and telephone-based machine inquiry systems enable users for information retrieval and transaction services. Such systems in the field of crisis management can be dedicated for providing appropriate information to ensure interoperability of emergency services and high-quality care for citizens. One challenge for such system is how to cope with an intense emotional and behavioral response of the users, such as grief, sadness, fear, pain, anxiety and anger, resulting from stressful conditions. Such emotions may cause diminished cognitive functioning, such as short term memory loss, confusion, difficulty setting priorities and making decisions (Farberow and Frederick, 1978), which can impair the ability to make sound decisions and take necessary steps toward resolving the crisis. The need of an automated dialogue system for crisis management that is able to perceive and respond user emotions becomes more apparent. How well these emotions are understood can impact the way the system can improve the quality of service or change the strategy of action selections or information presentation.

A dialogue system of a crisis management is dealing with several sources of uncertainty. Besides dealing with the errors in the speech recognition, the user can not be expected to always produce clean and correct grammatical sentences. When people speak spontaneously, they tend to produce a lot of extra sounds ("ehh", "umm"), correct themselves halfway a sentence, leave out parts of the sentence, and refer to concepts in a shortened and implicit way. A suitable compromise in such situations is to focus on the use of keywords in a sentence by shallow parsing. The input sentences can be analyzed by identifying certain keywords and constituents (such as, noun phrase, verb phrase, etc) without the need to specify their internal structure nor semantic.

Emotions can be expressed through our verbal and nonverbal behavior synchronously. We show the verbal emotion by our choice of words. Some words possess emotive meaning together with their descriptive meaning. The emotive meaning determines the "effects" of words, especially the emotional tone effect on the interpretation of one's speech contents. Such effect is a meaningful marker and an occasional mediator of our mental, social, and even physical state (Clore, 1992). Besides informing one's emotional state, the use of such words is also the bridge to reality (Ricoeur, 1976). The way we describe events can define the meanings of the events, which help us to improve our context awareness.

A crisis management is potentially overloaded with flood of inputs. Non-urgent and irrelevant reports, complains, and unnecessary questions, e.g. "my window is broken", "people are running away in panic", "I am afraid" and "is it over yet?", cannot help crisis responders seeking information about what is going on. However, information such as "we've just hit badly by a threatening earthquake", "I heard a nasty loud bang", and "a man is injured seriously" can describe the crisis event and hint the urgency level of the situations. A filter is necessary to distinguish non-urgent cases from urgent and emergent cases, which calls another requirement for the dialogue system.

The research reported here is part of a project running on developing a multimodal information presentation system applied in a crisis management. We develop a dialogue system that plays the role of a crisis hotline dispatcher receiving reports from people involved in a crisis event. It receives user inputs as text from speech-recognition (ASR) results. The system is able to reply users with appropriate prompts. An emotion analyzer from text has been deployed in this system, which can be used to analyze the user input emotional tone. Here, we explore a method on the way in which people use certain keywords to describe events, which marks their emotional state and (at the same time) indicates the urgency of their information.

The structure of this paper is as follows. In the following section, we start with related work. Further, our experimental dialogue system and our developed text-based emotion analyzer are presented in two sections, respectively. The performance of the system is indicated with some examples. Finally, we conclude the paper.

RELATED WORK

In developing a dialogue system, we must model the user's goal, the results retrieved, and the state of the dialog, and generate the system's response at each turn of the dialog. Many approaches to dialogue systems have been proposed as in (McTear, 2002; Catizone et al., 2002). To select a specific action in a dialogue flow, one of the approaches is by using a set of dialogue strategies, which can be formed in slots and heuristic rules (Rothkrantz et al., 2000), tables (Zue et al., 2000), frames (Catizone et al. 2003), task hierarchies (Potamianos et al. 2000). These systems basically use a set of rules and a state machine to analyze the structure of a dialogue and select a specific strategy. Colby's PARRY (1974) is known as the oldest dialogue system that is able to respond with emotions. Recent work in developing dialogue systems with emotions mostly is in the area of Embodied Conversational Agent (ECA, Cassell et al., 2000). The findings essentially address the impact of the user emotion state on the system's dialogue strategy and service performance.

Emotional linguistic content consists of entities of complexity and ambiguity such as syntax, semantics and emotions. Currently, text-based emotion analysis is approached mostly as a text-classification problem. A textual unit of certain size is classified as expressing positive or negative (or pleasant and unpleasant) feelings. The unit size can go from words (Hatzivassiloglou and McKeown, 1997; Hatzivassiloglou and Wiebe, 2000; Wiebe, 2000; Turney and Littman, 2003) to full texts (of various size), starting with a small set of seed words (Turney, 2002; Pang and Lee, 2004), manually built lexicons (Subasic and Huettner, 2001; Das and Chen, 2001), a mixture of unigrams, word sentiment measure, topic knowledge (Mullen and Collier, 2004), or even the world knowledge (Liu et al., 2003). Analyzing emotional aspect of language needs a large-scale affective lexicon resource database for example adjective database (Hatzivassiloglou and McKeown, 1997), common sense database (Liu et al., 2003), emotion class-based database (Ortony et al., 1978), affective-WordNet (Strapparava et al., 2004), emotion expression database (Desmet, 2002), and adjective database depicted based on their synonymy steps to "good" and "bad" (Fitrianie and Rothkrantz, 2006). The largest database, we found so far, is created by Whissell (1986) - DAL, which contains 8742 words in a 2D circumplex model.

CRISIS HOTLINE DISPATCHER

Our developed dialogue system applies human-human strategies in dialogue management for information services in the field of crisis management. Its dialogue manager consists of several modules (see figure 1). When the system receives input from a user, the dialogue manager interprets the user's utterances in the context of the ongoing dialogue: (a) the *input analyzer* interprets user's intention and retrieves information from the user's input and (b) the *emotion analyzer* analyzes the user's emotional state. This interpretation is then used to select the next action by the *dialogue act selector*. Information on the current state of the dialogue, domain knowledge and the user's emotional state are used to generate a response by the *response generator*.

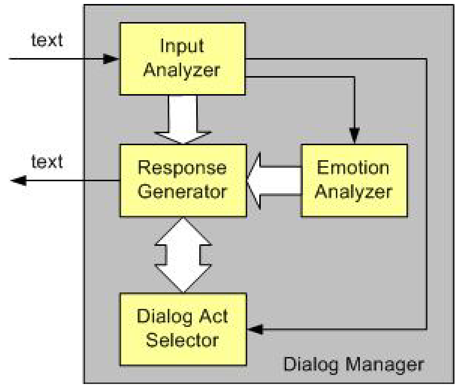

Figure 1 Architecture of the developed dialogue system

We developed our system's responses using an extended-XML script, AIML (Artificial Intelligent Mark-up Language - Wallace, 1995). AIML provides specifications for pattern input matching and reply generation. Besides matching user input patterns, its transformation rules allow user-system conversation focusing on a certain topic and generating responses based on previous conversation. The most important AIML elements are:

- <aiml>, the tag that begins and ends an AIML document.
- <category>, the tag that marks a "unit of knowledge" in a dialogue.
- <pattern>, the tag that contains a simple input pattern rule that matches what a user may type.
- <topic>, the tag that contains current conversation topic pattern rule.
- <that>, the tag that refers to the dispatcher's previous reply as a history pattern rule.
- <template>, the tag that contains the response to a user input.

We modified the AIML schema and added two new tags: (1) "`<frame>`", the tag that marks current dialogue action and (2) "`<concern>`", the tag that marks user's emotion state (Figure 2). Each category provides some possible response templates that can be selected depending on the value of both tags.

Each component in the developed dialogue manager is explained below.

Input analyzer

Our discourse model is developed based on (Rothkrantz et al., 2000). The model is constructed using goal-directed principles based on a crisis hotline dispatcher's point of view. The dispatcher bases his/her behavior on a comparison of a representation of the goal state and the current state. In Figure 3, the dispatcher's mental state during a dialogue is displayed. Based on Citizen Guide to 9-1-1 (2007), we designed the dispatcher's thoughts as seven frames (Table 1). Each frame consists of one or more slots. The goal of the dialogue system is to fill the frames and their slots.

All frames are empty at the beginning of the dialogue. Every state in the dialogue represents the dispatcher's thoughts at that time. A new state occurs after a dispatcher prompt and the user's reaction to it. Here, the *input analyzer* of our developed dialogue manager receives user text input from the ASR. It selects categories by ensuring that the most specific topic and previous reply match first before any of other categories or default (indicated by asterisk "*"). Then, the analyzer compares the input pattern by ensuring that the most specific pattern matches first. Inside the template, some set tags are used to fill slots' value in a certain frame:

```
<set name="frame.<frame-name>.<slot-name>" value="<value>"
     type="["filled", "assume", "uncertain", "question"]"
\>
```

and to set the current dialogue topic "`<set name="topic" …>`". The analyzer decides the value's type whether it is extracted from the input ("filled"), assumed by the dispatcher ("assume"), or uncertain ("uncertain"). If the slot is asked by the user, the type is "question" and the value is empty. Based on these set tags, the input analyzer fills the value of the defined slots. For example, when the dispatcher receives "A building in front of my restaurant is collapsed”, the dispatcher sets frame.caller.type’s value = “WITNESS” and type = “filled” (Figure 2). Figure 3 shows that the caller frame becomes semi-filled.

```
…
<topic name=”*”>
  <category>
    <pattern>* IS COLLAPSED</pattern>
    <that>HOW</that>
    <template>
       <set name=”frame.problem.type” type=”uncertain” value=”EXPLOSION”/>
       <set name=”frame.problem.status” type=”assume” value=”DANGER”/>
       <set name=”frame.reason.type” type=”uncertain” value=”TERRORIST-ATTACK”/>
       <set name=”frame.urgency.status” type=”assume” value=”HIGH”/>
       <set name=”frame.caller.type” type=”filled” value=”WITNESS”/>
       <set name=”frame.time.*” type=”assume” value=”system.getDateTime()”/>
       <set name=”topic” value=”EXPLOSION”/>

       <frame name=”location.*”>
          <concern valence=”T<=-0.5” arousal=”T>= 0.3” aggressive=”FALSE”>
           I’ll pass this to the police department. Could you give me your address?
          </concern>
          <concern valence =”T<=-0.5” arousal=”T>=0.5” aggressive=”TRUE”>
           We are in difficult situation. Please tell me your
           address so the police can come to check </star>
          </concern>
          … // other concern
       </frame>
       <frame name=”involved_party.status”>
           <concern valence=”-0.6<T<0” arousal=”-0.2<=T<=0.2” aggressive=”FALSE”>
            Do you see any victim?
           </concern>
           … // other concern
       </frame>
       <frame name=”reason.*”>
          <concern valence=”-0.6<T<0” arousal=”-0.2<T<0.2” aggressive=”FALSE”>
           What do you mean? Is there any explosion?
          </concern>
          <concern valence=”0.2<T<0” arousal=”T<-0.2” aggressive=”HIGH”>
           Tell me what happen.
          </concern>
          … // other concern
       </frame>
       <frame name=”*”>
          <concern valence=”T>=0” arousal=”T<-0.2” aggressive=”*”>
           All emergency services are busy answering helps. Please hold on until you
           learn if there is any casualty.
          </concern>
          … //other concern
       </frame>
       … //other frame
    </template>
  </category>
…
</topic>
… // other topic
```

Figure 2 An example unit in the AIML database; asterisk (*) means “any”

Table 1 Dialogue frames and their slots

Frame	Slot
caller	{name, address(street, no, city, postcode), phone-no, role[“VICTIM”, “WITNESS”, “PARAMEDIC”, “FIREMEN”, …], valence-score, arousal-score, aggressive-sign}
problem	{type[“FIRE”,”TRAP”,’SICK”, …], status, desc}
time	{hour, minute, second, day, month, year}
location	{street, no, city, postcode}

Frame	Slot
`involved_party`	`{type["VICTIM", 'SUSPECT", "PARAMEDIC", …], status, address(street, no, city, postcode), name, phone-no}`
`reason`	`{type["HIT", "EXPLOSION","GAS-LEAKING", …]}`
`weapon`	`{type["GUN", "KNIFE","BOMB",…]}`
`urgency`	`{type[HIGH, MEDIUM, LOW]}`

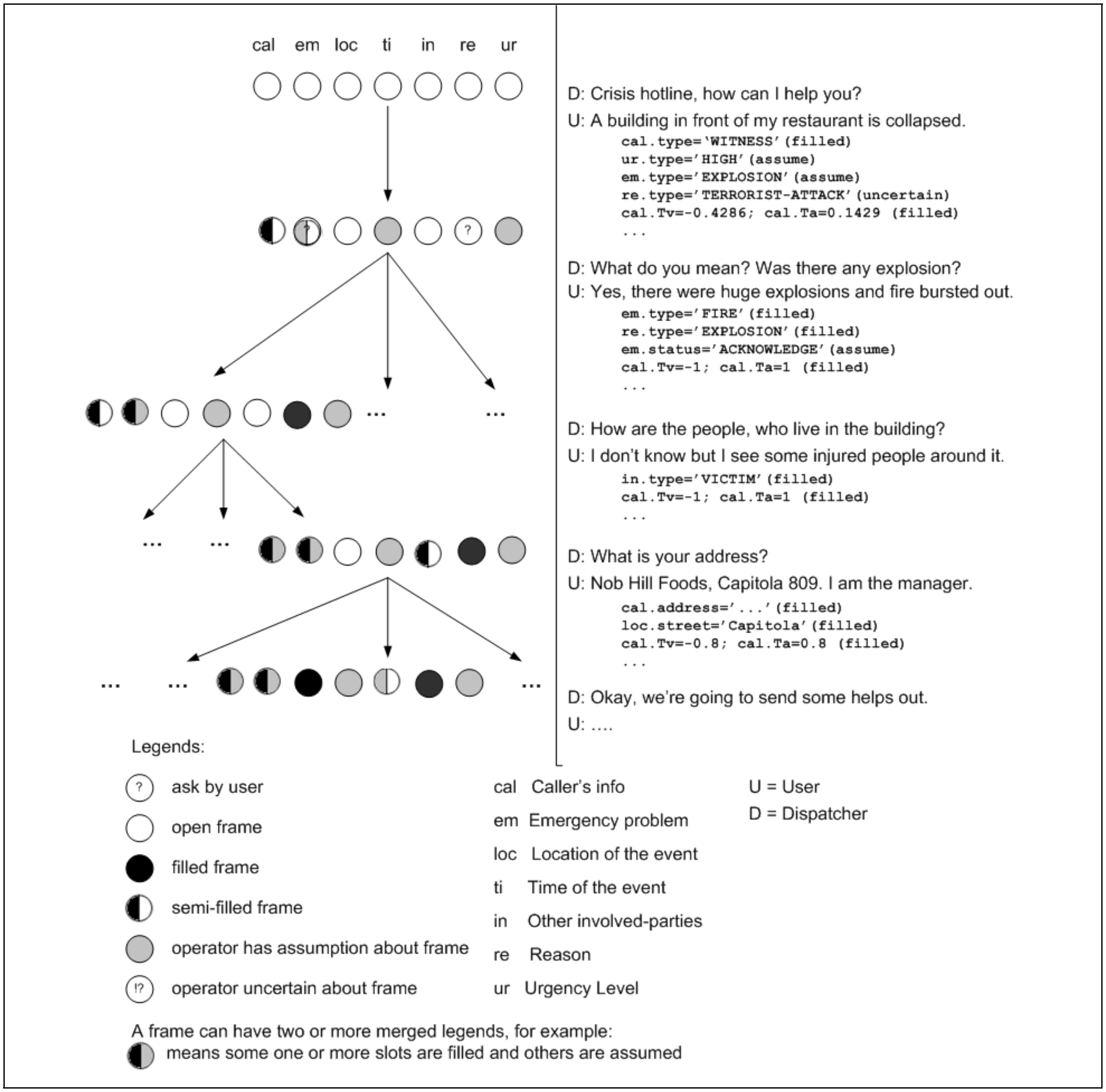

Figure 3 Dialogue model from dispatcher perspective (circles represents frames)

Emotion Analyzer

At every state of user-system dialogue, there are different ways to continue a dialogue. One way is based on user's current emotional state, which is indicated by the concern tag. The *emotion analyzer* has a parser for analyzing emotion loading on text. Figure 4 shows our developed method, which consists of two steps: (a) processing phase and (b) analysis phase. The *processing phase* extracts the input sentence based on its constituents. It uses a parser to select only those that contain certain keywords and aggressive words. The *analysis phase* calculates the valence and arousal scores of these constituents. The valence and arousal scores of an input are the average scores of all selected constituents in the input. The value of two thermometers T of valence and arousal will then changes based on this calculation (see next section for detailed calculation). Besides for selecting the next action, these values are used to fill in caller-info frame's slots. This analyzer also returns an urgency value $u_d(v,a)$ of the information that is given by the user.

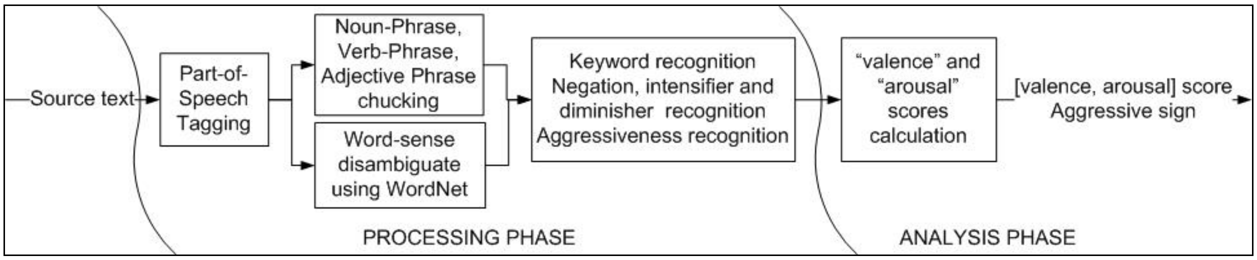

Figure 4 Schematic view of emotion recognition from text

Dialogue Act Selector

Another way to select the next action of user-system dialogue is based on an evaluation of how much dialogue steps will reduce the distance to the goal state, which implies that in the successor action a maximal slot-filling is achieved. For this purpose, the current implementation of our *dialogue act selector* uses a heuristics rule-based approach to choose the next action based on the nature of emergency problem. Firstly, we indicated optional frames and slots based on the emergency problem. For example, in the case of the "`problem.type`" is "FIRE" some optional slots such as `involved_party.name` and `problem.desc` may not be filled in. Finally, we designed the heuristics rules, for example:

- if there is a (non-empty) subset of open frames and a (non-empty) subset of filled frame, ask for information covering as much as the open frames as possible;
- if one frame is in the current state is not completely filled, immediately ask for the missing information to solve the ambiguity concerning this frame;
- if the subset of not complete filled frames contains more that one slot, handle the individual slots one after the other;
- as long as new information can be provided assumptions of the dispatcher are not verified.

Figure 3 shows the differences in the dispatcher's thoughts in every dialogue state as compared to the previous state. The output of this module is a set of possible (obligatory) frames and slots to be filled in the next dialogue state.

Response Generator

The task of the *response generator* is to generate the actual system response. Given a set of possible (open) frames and slots to choose, this module selects an associated frame's name and slot's name,

```
<frame name="<frame-name>.<slot-name>">
```

from the templates. The most specific frame's and slot's name is matched first before any frames or slots. Within the selected frame tag, given the user emotional state (the value of $T_{valence}$ and $T_{arousal}$), the module selects the most appropriate response based on concern tag,

```
<concern valence ="<valence-degree>" arousal="<arousal-degree>"
          aggressive="[TRUE, FALSE]">
```

Figure 7 shows examples of a dialogue fragment between user and our crisis online.

Conversation 1:	
Dispatcher:	Crisis hotline. What happened?
User:	Is it over yet?
Dispatcher	I do not have such information.
User:	Oh fuck! .. I am in Bay Reeds.
Dispatcher:	Calm down please. Anyone there hurt?
User:	I am not sure … Damn! ..
Dispatcher:	Tell me what you know. Please calm your self.
User:	I just cannot take it anymore.
Dispatcher:	Hanging there, okay?
User:	Okay.
Dispatcher:	If you know if there is any injury or smell of gas, or fire, call us back please. Bye.
…	**$u_d(v,a)$= (0,0), T_v=-0.4, T_a=0.4, aggressiveness = TRUE**
Conversation 2:	
Dispatcher:	Dispatch Center. What are you reporting?
User:	I heard great nasty explosions coming from the north of my house, followed by earthshaking.
Dispatcher:	What kind of explosion? Any injury?
User:	It must be a big bomb. I can't see anything from my house.
Dispatcher:	Where is your address? Please stay in your house.
User:	220 Forth North.
Dispatcher:	Thank you.
User:	Okay.
Dispatcher:	If you learn any casualties, please call us back.
…	**$u_d(v,a)$= (-1,1), T_v=-0.6, T_a=0.6, aggressiveness = FALSE**

Figure 5 Dialogue fragments between a user and our developed crisis hotline-bot. Information from the second dialogue is much more urgent than from the first one

AUTOMATED TEXT-BASED EMOTION ANALYZER

Our emotion analyzer relies on indicators on the text input, such as nouns, verbs, adjectives and adverbs, and uses semantic scores provided by the DAL (Whissell, 1986). Since semantic orientation may diminish or even change (Kennedy and Inkpen, 2006), in the task of automatically classifying the input as positive or negative emotion, the analyzer considers three types of emotional shifters: negations (they reverse the polarity of a term), intensifiers and diminishers (they affect the degree to which a term is positive or negative).

Two-Dimensional Affective Lexicon Dictionary

This emotion analyzer uses four dictionaries, such as:

1. Collection of nouns, verbs, adjectives and adverbs from DAL. Each word has a valence and arousal score in the interval [1 .. 3] (Figure 5), e.g. the valence score of “bad” is 1.2857 and the arousal score is 1.4615.
2. Collection of 200 keywords in the field of crisis (103 verbs, 27 adjectives and 70 nouns) from 12 news articles with six different topics (such as fire, bombing, flooding, tsunami, earthquake, and terrorist attack) and four 9-1-1 recording transcriptions (SFMuseum, 1989, WNBC, 2002). More than 92% are already known in DAL.
3. Collection of 85 aggressive-sign words from the Alternative Dictionary. Some words are known in DAL.
4. Collection of 28 qualifiers (“very”, “rather”, “huge”, “low”) and 32 compound qualifiers (e.g. “not very low”, “more or less high”, and ‘sort of high”), which are ranked heuristically based on their intensity I (from “not at all” to “very” and from “tiny” to “huge”) and are assigned to a value $I_q \in [0..1]$ for each word. For example $I_{\text{very}} = 1$, $I_{\text{medium}} = 0.5$, $I_{\text{slightly}} = 0.2$, $I_{\text{not very low=0.15}}$ and $I_{\text{not at all}} = 0$.

For those words in that are out of DAL's domain, we used WordNet (Fellbaum, 1998) to assign their valence and arousal scores based on their synonym and hypernym.

Figure 6 38 adjectives from DAL depicted in 2D "valence" and "arousal" in the interval [-1..1]

Relevant Information Selection

The *processing phase* uses the QTAG POS Tagger (Tufis and Mason, 1998), which returns the part-of-speech of each word (e.g. noun, verb, and adjective) in a sentence based on a (n-grams) probabilistic tagger. A parser extracts the output into a collection $S = \{p_1, p_2, \ldots p_n\}$ where p can be noun phrases, verb phrases (in particular, a set of verb(s) and adverb phrase(s)) or adjective phrases (a set of noun phrase(s) and adjective(s)). Each phrase is a tupple $p_i = (m_i, e_i)$, where m_i can be one or more verbs or nouns and e_i is a set of words or phrases that explains m_i. Each e_{ij} is defined as $e_{ij} = (q_{ij}, \tau_{ij})$, where q_{ij} is one or more qualifier that can intensify or diminish the absolute quality of the fact τ_{ij}. For example: "we're just attacked very badly by a quite threatening and long earthquake" is extracted into p_1 = ("we", null), p_2 = ("are attacked", ("very", "badly")) and p_3 = ("earthquake", (("quite", "threatening"), (null, "long")). The parser discards those phrases that do not contain any crisis keyword. Therefore from the previous example, we have $S = \{p_2, p_3\}$ since "attack" and "earthquake" are keywords in the crisis field. In addition, this parser also collects aggressive words.

For any negation word (e.g. "not", "no", "never"), the parser executes the following algorithm:

```
if the negation negates any mi in a certain pi then
   this pi is discarded from S
else if the negation negates any τij in a certain eij then
   this τij will be replaced by its direct-antonym (based on WordNet) in the DAL
else the negation is treated as qualifier
```

For example: the phrase "(is) not an earthquake" will be discarded from the collection because its affective orientation is null. The phrase "not a severe earthquake" will be replaced by "a simple earthquake". The phrase "not very" in "not very long earthquake" will be treated as a qualifier and has an intensity value $I_{\text{not-very}}$. The parser assigns valence and arousal scores for all words.

Word-Orientation Count Method

From the processing phase, for each input source S, we have:

$$S = \{p_1, p_2, \ldots, p_n\} \mid p_i = \{m_i, e_i\},\ m_i = \{w_{i1}, w_{i2}, \ldots, w_{in}\},\ e_i = \{(q_{i1}, \tau_{i1}), (q_{i2}, \tau_{i2}), \ldots, (q_{in}, \tau_{in})\}$$

where w_i can be a noun or verb. In the *analysis phase*, our developed emotion analyzer calculates the valence and arousal scores of S in the following five steps.

1. Convert the valence and arousal scores into the interval [0..1].

   We consider that the valence score of a word corresponds to the probability of this word being pleasant. For example, the value of $v_{word} = 2$ means: (a) the word is neutral and (b) v_{word} is equivalent with probability value

0.5 since 50% raters rates the word being positive, which implies that with the same probability the word is negative. This also works for the arousal score, which corresponds to the probability of a word being active. Therefore:

$$\forall word_i \; in \, DAL \,|\, v'_{word_i} \in [0..1] = (v_{word_i} - 1)/2$$
$$\forall word_i \; in \, DAL \,|\, a'_{word_i} \in [0..1] = (a_{word_i} - 1)/2$$

2. Calculate the valence and arousal scores of m_i.

We consider that the valence and arousal scores of m_i are calculated as the extreme scores of all w_{ij} from the largest number of w_{ij} with the same emotional orientation. For example, if more than 50% of w_{ij} has *score>0.5*, the maximum score will be taken.

$$v_{m_{ij}} = EXTREME\_VALUE(v_{w_{i1}}, v_{w_{i2}}, ..., v_{w_{in}})$$
$$a_{m_{ij}} = EXTREME\_VALUE(a_{w_{i1}}, a_{w_{i2}}, ..., a_{w_{in}})$$

$$S(x;\alpha,\beta,\gamma) = \begin{cases} 0 & for \;\; x \le \alpha \\ 2\left(\dfrac{x-\alpha}{\gamma-\alpha}\right)^2 & for \;\; \alpha \le x \le \beta \\ 1-2\left(\dfrac{x-\gamma}{\gamma-\alpha}\right)^2 & for \;\; \beta \le x \le \gamma \\ 1 & for \;\; x \ge \gamma \end{cases}$$

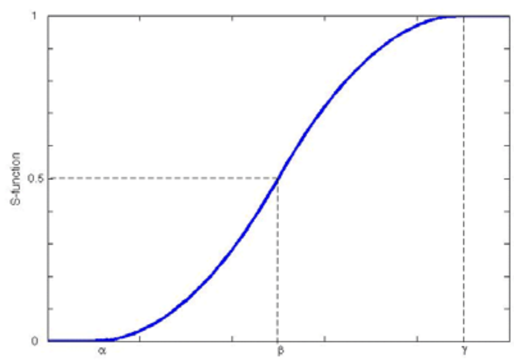

Figure 7 S-Function(x; α, β, γ)

3. Calculate the valence and arousal scores of e_i.

Since the quality of τ_{ij} can be diminished or intensified by the existing q_{ij} within a statement, all these qualifiers are represented and operated in fuzzy theory. We consider that the I values ($\in$[0..1]) forms a fuzzy set of q_{ij}. We used the inverse of S-function (Figure 6) to model the membership function of this fuzzy set, which treated the valence and arousal scores as parameters to adjust the membership function. We assume that the new valence and the arousal scores of τ_{ij} will never be less than 0, since the negation of a term has been taken care in the processing phase.

$$I_{q_{ij}} = S(v_{\tau_{ij}}; v_{\tau_{ij}} - \xi, v_{\tau_{ij}}, v_{\tau_{ij}} + \xi) = S(a_{\tau_{ij}}; a_{\tau_{ij}} - \xi, a_{\tau_{ij}}, a_{\tau_{ij}} + \xi)$$
$$v'_{\tau_{ij}}(I_{q_{ij}}, v_{\tau_{ij}}, \xi) = S^{-1}(v_{\tau_{ij}}; v_{\tau_{ij}} - \xi, v_{\tau_{ij}}, v_{\tau_{ij}} + \xi) \,|\, v'_{\tau_{ij}} = 0 \; if \; v'_{\tau_{ij}} \le v_{\tau_{ij}} - \xi \; and \; v'_{\tau_{ij}} = 1 \; if \; v'_{\tau_{ij}} \ge v_{\tau_{ij}} + \xi$$
$$a'_{\tau_{ij}}(I_{q_{ij}}, a_{\tau_{ij}}, \xi) = S^{-1}(a_{\tau_{ij}}; a_{\tau_{ij}} - \xi, a_{\tau_{ij}}, a_{\tau_{ij}} + \xi) \,|\, a'_{\tau_{ij}} = 0 \; if \; a'_{\tau_{ij}} \le a_{\tau_{ij}} - \xi \; and \; a'_{\tau_{ij}} = 1 \; if \; a'_{\tau_{ij}} \ge a_{\tau_{ij}} + \xi$$
$$where \;\; \xi = x - \alpha = \beta - x \;\; as \, an \, adjusted \, coefficient \, factor$$

Here, we took three assumptions: (1) $a_{\tau_{ij}}$ and $v_{\tau_{ij}}$ values always correspondence to I =0.5, (2) ξ =0.5, and (3) if τ_{ij} is null, e.g. p_1= {“earthquakes”, (“small”, null)} with I_{small}, we assume that τ_{ij} = m_ij. The valence and arousal scores of e_i are calculated as the extreme new scores of all τ_{ij} from the largest number of τ_{ij} with the same emotional orientation.

$$v_{e_i} = EXTREME\_VALUE(v'_{\tau_{i1}}, v'_{\tau_{i2}}, ..., v'_{\tau_{in}})$$
$$a_{e_i} = EXTREME\_VALUE(a'_{\tau_{i1}}, a'_{\tau_{i2}}, ..., a'_{\tau_{in}})$$

4. Calculate the valence and arousal scores of p_i.

   Given the independent m_i and e_i being pleasant with probability v_{m_i} and v_{e_i} and being active with probability a_{m_i} and a_{e_i}, respectively. The valence score of phrase p_i is equal to the combination of m_i and e_i, which is considered as the probability of union m_i and e_i,

   $$v_{p_i} = (v_{m_i} + v_{e_i}) - (v_{m_i} \bullet v_{e_i})$$
   $$a_{p_i} = (a_{m_i} + a_{e_i}) - (a_{m_i} \bullet a_{e_i})$$

   This calculation can be recursive to calculate v_e and v_m. For example, for a compound sentence "the earthquake, which shook the earth severely", the system will calculate p_1= {"shook", (null, "severely")} and then p_2= {"the earthquake", (null, $p1$)}.

5. Calculate the valence and arousal scores of S.

   These scores are the mean of all collected phrases in S.

   $$v_s = \frac{1}{n}(v_{p_1}, v_{p_2}, ..., v_{p_n})$$
   $$a_s = \frac{1}{n}(a_{p_1}, a_{p_2}, ..., a_{p_n})$$

User Valence and Arousal Thermometer

Two emotion thermometers, such as valence T_v and arousal thermometer T_a, are used to observe the intensity of the user emotional state during conversation. Here, we choose to use the interval [-1..1] to give natural appealing of emotional and activation orientation. Therefore:

$$v'_s = 2v_s - 1 \mid v'_s \in [-1..1]$$
$$a'_s = 2a_s - 1 \mid a'_s \in [-1..1]$$

For every new user input, this emotion analyzer module will calculate all thermometers using the following equations:

$$T_v(t+1) = \begin{cases} 1, T_v(t) + v'_s \geq 1 \\ T_v(t) + v'_s + \varphi \\ -1, T_v(t) + v'_s \leq -1 \end{cases}, \quad T_a(t+1) = \begin{cases} 1, T_a(t) + a'_s \geq 1 \\ T_a(t) + a'_s - \varphi \\ -1, T_a(t) + a'_s \leq -1 \end{cases}$$

The aggressive sign is defined as Boolean TRUE or FALSE, whether or not the user uses any aggressive words in the text input. φ is the neutral factor of the user's emotional state, when `aggressiveness = FALSE` and $v'_s = a'_s = 0$. If `aggressiveness = TRUE` and $v'_s = a'_s = 0$, the φ becomes negative.

The value of both thermometers is considered as the user's current valence and arousal scores in the interval [-1..1]. Table 2 shows some examples of sentences given by users.

Urgency Level of User Information

The urgency value $u_d(v,a)$ of information is defined as the minimum valence score and the maximum arousal score of all source inputs from a dialogue d.

$$u_d(v,a) = (MIN(v'_{s_1}, v'_{s_2}, ..., v'_{s_n}), MAX(a'_{s_1}, a'_{s_2}, ..., a'_{s_n}))$$

This value is classified as follows:

$$urgency\deg ree = \begin{cases} HIGH \Leftrightarrow u_d.v \leq -0.2\, and\, u_d.a \neq 0 \\ LOW \Leftrightarrow -0.2 < u_d.v < 0 \\ NEUTRAL, otherwise \end{cases}$$

This value is used to fill a slot in the urgency frame.

Table 2 Some examples of user inputs. The negative orientation: D>C>B, A>B, C~A

No.	Input	Valence score	Arousal score
A	I saw a building collapsing.	-0.4286	0.1429
	Fire burst out.	-0.6364	0.667
	I just heard an explosion.	-0.778	1
B	A little fire occurs around the building .	-0.3504	0.6098
	Small explosions near the neighborhood.	-0.5871	0.8134
C	There's more less big fire at the Conference Center.	-0.6241	0.6616
	The explosion is somewhat terrifying.	-0.7753	0.9922
D	An extremely terrible fire burns the building.	-0.9285	0.8335
	I heard great nasty explosions.	-1	1

CONCLUSION AND DISCUSSION

Our experimental dialogue system of a crisis management is controlled by a frame-based computational model of the dispatcher thoughts about the hypotheses and the user's intentions. The model applies rules and strategies for selecting actions to extract information from user side, which causes the dispatcher's behavior to be rather slot-oriented. The next development will focus on adding a capability on accessing external sources (e.g. a database) to respond the user's query. We develop a text-based emotion recognition that considers emotional shifters, such as negations, intensifiers and diminishers, in text input. Future work still needs to be done to study more corpora of human-human dialogues and to evaluate the heuristics approach of the emotion recognition with different sentence structures and human users. An ongoing work is in adding AIML-dialogue units with more topics based on the nature of specific emergency problems. In particular, the system's responses are designed to have more open questions. Besides to achieve maximal (optimum) slot-fillings, with a large number of possible responses, we expect that users will have natural conversation with the developed system and the emotion recognition module will result more accurate analysis. We assume that choosing an appropriate and useful response of the dispatcher can provide new information, reduce the uncertainty and provide clearness, whereas an inappropriate response takes extra time and may cause confusion. Retrieved information from the users can, then, be used for creating situation awareness or supporting decision making. The primary results show that our text-based emotion analysis approach offers a simple method for user's emotional state analysis to investigate. However, its accuracy level depends on the accuracy level of the affective lexicon database. Currently, we develop emotion recognition from speech prosody. We found it quite difficult to find human-human dialogue recordings that are free from noise. However, we expect that the result of our text-based emotion analyzer can enhance and complement the result of the speech-based emotion analyzer.

ACKNOWLEDGMENTS

The research reported here is part of the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024. We like to thank Alin G. Chițu for his contribution.

REFERENCES

1. Cassell J., Sullivan J., Prevost S. and Churchill E. (2000). *Embodied Conversational Agents*, MIT Press Cambridge.
2. Catizone R., Setzer A. and Wilks Y. (2002) *State of the Art in Dialogue Management*. In: Deliverable D5.1 of COMIC Project, http://www.hcrc.ed.ac.uk/comic/documents/.
3. Catizone R., Setzer A. and Wilks Y. (2003) Multimodal Dialogue Management in the COMIC Project, *EACL*, Hungary.
4. Clore G. (1992) Cognitive Phenomenology: Feelings and the Construction of Judgment, In: L. Martin and A. Tesser (Eds.), *The Construction of Social Judgments*, NY: Lawrence Erlbaum Associates, 133-163.

5. Colby K. M., Parkinson R.C. and Faught B. (1974) Pattern-Matching Rules for the Recognition of Natural Language Dialogue Expressions, *ACM Linguistics*.
6. Das S. and Chen M. (2001) Yahoo! For Amazon: Sentiment Parsing from Small Talk on the Web, *APFA*, Thailand.
7. Desmet P. (2002) *Designing Emotion*, Doctoral Dissertation, Delft University of Technology.
8. Ekman P. (1999) Basic Emotions, In Dalgleish T. and Power M., (Eds). *Handbook of Cognition and Emotion*, UK: John Wiley and Sons, Ltd.
9. Farberow N.L. and Frederick C.J. (1978) *Training Manual for Human Service Workers in Major Disasters*. Rockville, Maryland: National Institute of Mental Health.
10. Fellbaum C. (1998) *WordNet: An Electronic Lexical Database*. The MIT Press.
11. Fitrianie S. and Rothkrantz L.J.M. (2006) Constructing Knowledge for Automated Text-Based Emotion Expressions, *CompSysTech 2006*, Bulgaria.
12. Hatzivassiloglou V. and Mckeown K. (1997) Predicting the Semantic Orientation of Adjectives. *ACL*, Spain, 174–181.
13. Hatzivassiloglou V. and Wiebe J. (2000) Effects of Adjective Orientation and Gradability on Sentence Subjectivity, *COLING*, Germany, 299–305.
14. Kennedy A. and Inkpen D. (2006) Sentiment Classification of Movie Reviews using Contextual Valence Shifters. *Computational Intelligent,* 22 (2):110-125.
15. Liu H., Lieberman H. and Selker T. (2003). A Model of Textual Affect Sensing Using Real-World Knowledge. *IUI*, Florida, 125–132.
16. McTear M.F. (2002) Spoken Dialogue Technology: Enabling the Conversational User Interface. *ACM Computing Surveys*, 34:90–169.
17. Mullen T. and Collier N. (2004) Sentiment Analysis using Support Vector Machines with Diverse Information Sources. *EMNLP*, Spain, 412–418.
18. Pang B. and Lee L. (2004) A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. *ACL*, Spain, 271–278.
19. Potamianos A., Ammicht E. and Kuo H-K. (2004) Dialogue Management in the Bell Labs Communicator System. *ICSLP*, China.
20. Ortony A., Clore G. and Collins A. (1988) *The Cognitive Structure of Emotions*, Cambridge University Press.
21. Ortony A., Clore G. and Foss M. (1987) The Referential Structure of the Affective Lexicon, *Cognitive Science*, 11: 341-364.
22. Ricoeur P. (1976) *Interpretation Theory: Discourse and the Surplus of Meaning*. Fort Worth, TX: Texas Christian Univ. Press
23. Rothkrantz L.J.M., van Vark R.J., Peters A. and Andeweg N.A. (2000) Dialogue Control in the Alparon System, *TSD,* in P.Sojka, I. Kopecek and K. Pala (Eds.), LNAI 1902, 333-338.
24. Russell J.A. (1980) A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, 39(6) 1161-1178.
25. Subasic P. and Huettner A. (2001) Affect Analysis of Text Using Fuzzy Semantic Typing. *IEEE-FS*, 9:483–496.
26. Strapparava C. and Valitutti A. (2004) WordNet-Affect: An Affective Extension of WordNet, *LREC*, Portugal.
27. The Alternative Dictionary. (1994-2002) *Slang, Profanities, Insults and Vulgarisms from All the World*, http://www.notam02.no/%7Ehcholm/altlang/.
28. The Virtual Museum of the City of San Francisco. (1989) *San Fransisco 9-1-1 Dispatch Tapes*, 17 October, http://www.sfmuseum.org/1989/sf911.html and http://www.sfmuseum.net/1989/sc911.html.
29. Turney P. (2002) Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification Of Reviews. *ACL*, Philadelphia, 417–424.
30. Turney P. and Littman M. (2003) Measuring Praise and Criticism: Inference of Semantic Orientation from Association. *ACM Transactions on Information Systems*, 21(4):315–346.
31. Wiebe J. (2000) Learning Subjective Adjectives from Corpora. *AI and Conf. on IAAI*, Texas, 735–740.
32. Whissell C. M. (1989). The Dictionary of Affect and Language. In Plutchik, R. and Kellerman, H. (eds), *Emotion: Theory, Research, and Experience*. Academic Press, New-York, 4: 113-131.
33. WNBC. (2002) *Exclusive: 911 Tapes Tell Horror of 9/11 (Part 1 and 2): Tapes Released for First Time*, http://www.wnbc.com/news/1315646/detail.html.
34. Zue V., Seneff S., Glass J., Polifroni J., Pao C., Hazen T. and Hetherington L. (2000) JUPITER: A Telephone-Based Conversational Interface for Weather Information. *IEEE Trans. Speech and Audio*, 20/Y, 100-112.