A Mobile Intelligent Dialogue Agent for gas classification in crisis situations

BSc Artificial Intelligence (*Bachelorscriptie*)

written by

Joris de Groot
0518808

under the supervision of **Dr. Vanessa Evers** and **Dr. Andi Winterboer**, in partial fulfilment of the requirements for the degree of

BSa Artificial Intelligence

at the *University of Amsterdam*.

June 26, 2009

Abstract

This thesis presents a Mobile Intelligent Dialogue Agent that can be used in the crisis management project DIADEM. The Agent is used in crisis situations where a chemical accident near a populated area has occurred and a gas is released. The Agent is implemented as a mobile application and can interview people on their mobile phone to gather information on what they smell.

Research is done on olfactory perception to see how people describe odors and a gas database is created where gases are described in more detail. A Bayesian Belief Network then is implemented to classify what people smell using the answers from the people. The Agent uses a simple interview style where only ”yes/no” questions are asked. This way of interviewing people to gather information is faster then the current method, which is over the phone. Next to this, the Agent uses the Bayesian Network to select the most probable questions given the probabilities of the nodes in the network. In this way, the interview can adapt to the current crisis situation and the interview will be as short as possible.

This has resulted in a first implementation of the Mobile Intelligent Dialogue Agent for the Android operating system. The Agent is able to select the most probable questions first and in theory can also adapt to the crisis situation.

Contents

1 Introduction

Mobile devices are common practice in today's world. Virtually everyone has a mobile phone and therefore it can be very useful in crisis situations. For instance, when a crisis situation occurs, people can be contacted over their mobile phone to inform them about the current situation. But people nearby the location of the crisis could also be contacted to help, by providing responders with valuable information. This thesis describes an application for the latter case: an intelligent mobile agent that can contact people over their mobile phone to extract information about the crisis situation. The crisis domain will consist of gas leaks in industrial area's which produce stench. The agent uses a Bayesian Believe Network (BBN) to classify what people smell by asking "yes/no" questions and then uses the answers to these questions to discover what gas causes the stench. By learning from the answers of the people, the BBN will be able to ask the more relevant questions earlier and in this way be able to identify the category of the smell faster. After determination of the most likely category, the agent will have a smaller number of candidate gases from which the most probable should be chosen. This is done by asking the questions which reduce the entropy (measure of uncertainty) for these candidate gases the most. The agent thus adapts to the current situation and to the answers of the people, which results in faster classification. In this way the fewest time from the user is taken and valuable information is collected fast, which is very important in a crisis situation.

The idea for this Intelligent Dialogue Agent comes from the DIADEM project (Pavlin, Wijngaards, & Nieuwenhuis, 2009). The DIADEM project is a collaborative project "which focuses on a novel combination of advanced technologies which facilitate collaborative information processing in environmental management applications". The DIADEM system will be used in industrial areas where there is a potential health hazard for people who live or work in these area's. When a potential dangerous situation arises, which is the case when there is a gas leak in a factory near a populated area, there is a need for a lot of different information. Information about the gas that is released, about the weather and the number of people potentially in danger must be received and processed by the appropriate experts. With this information, decision makers like the mayor can make the best informed decisions in these crisis situations. The DIADEM system will support the decision process by creating an infrastructure for all the different kinds of information and by automating some of the simple parts of the decision process. In this way, the information that is being received during a crisis situation will be send to the appropriate experts at the appropriate time. This saves the experts the task of collecting this information themselves, which gives them more time to focus on the situation at hand.

The Intelligent Dialogue Agent described in this thesis can be used at the start of collecting information for the DIADEM project. For instance, when a gas is released it is possible that the leak is not immediately noticed by the factory staff. The gas can then spread over a populated area which can cause stench. When people smell this stench they can contact the authorities [1], usually over the phone, which triggers an investigation to the source of the smell. The DIADEM system will then assist during such an investigation. Currently, this investigation is started after around five people have called to report some kind of anomaly. These calls take a lot of time to process since there are only a few operators present at the control room and people who call tend to elaborate a lot and drift away from the actual topic: describe what they smell. The Intelligent Dialogue Agent can make the process of gathering information from people much faster, since it can contact anyone with the this agent installed on their phone and stop people from elaborating too much by simplifying the interviews. Next to this, the agent can make a start with classifying the anomaly that is being smelled by a user. By using this simplified interview with questions on which people can only answer with "yes" or "no", more information about an anomaly can be assembled. With this information the agent can then classify the smell into categories. This start of the classification can then be used by the appropriate experts without anyone having to do extra work. This thus saves time and time is essential in crisis situations.

Finally, the agent will function as a bridge between the complicated and intelligent DIADEM system and the simple process of information extraction from people via mobile devices, while still assisting in the classification of gases. This way of extracting information saves time, because the staff of the

[1]For this project the authority is DCMR Milieudienst Rijnmond (www.dcmr.nl)

control room don't have to filter the usable information from the phone calls and they don't have to feed this information to a decision networks because this process is automated. When enough people would have this agent installed on their mobile phone and are willing to participate in these kinds of interviews with this agent, some statistically valid conclusion about what gas was released during the incident can be made. The conclusions of this agent can then be used by human experts to do further investigations with a decreased number of hypotheses and with the probabilities for these gases, which saves a lot of valuable time.

The remainder of this thesis will be organised as follows: in section 2 an overview of related work in the domains of crisis management, mobile adaptive systems and olfactory research is given. Then in section 3 the different problems that have to be solved to create this agent are defined. Then a description of the implementation of the agent is given in section 4. In section 5 the results can be found, followed by the limitations and future work in section 7 and finally the conclusions and discussion in section 8.

2 Overview of previous work

As stated in the introduction, this section will give an overview of previous work in the domains about the use of mobile agents in crisis management and about olfactory research. These categories will be discussed separately.

2.1 Mobile agents in crisis management

Crisis management is an area of research within different disciplines. As mentioned before, this project has the goal to develop a mobile dialogue agent that can be used in crisis situations where a chemical incident has occurred. An overview of the aspects of crisis management will follow next.

In (Reddy et al., 2009) an overview is given of the major challenges with coordination in crisis situations. This article gives a good overview of the important aspects of crisis management. The articles describes a research done on the communication and coordination between emergency medical services, like trauma helicopters and ambulances, and emergency department teams, like doctors and nurses in a hospital. Reddy *et al.* state that "one of the key factors for effective crisis management is designing information and communication technologies (ICTs) that support effective and seamless coordination between teams during a crisis". This is also one of the goals of the DIADEM project, where a lot of different experts need all kinds of different information that has to be communicated between them. Then the major challenges associated with team coordination in crisis management are said to be "information mismanagement, resource allocation issues and ineffective communication". Reddy *et al.* then do a research where focus groups are given a scenario where a leakage of hazardous materials has occurred. They observe how the different teams cooperate and information is coordinated within teams and between teams. From this research three major challenges could be identified namely *1*) ineffectiveness of current information and communication technologies *2*) lack of common ground *3*) and break downs in information flows

The first point illustrates that there is a need for new and better communications techniques that can communicate information better and more efficient. The second point says that there is a constant problem of maintaining a "mutual knowledge", so information that is known by everyone in all the teams. This is important in crisis situations, because when two people from different teams have to communicate it is important that you know what the other guy knows. In that way, no time is lost updating people who are communicating with the most recent information. The last point illustrates that it is important that all people who need to be informed or updated with the newest information are available. So when a new peace of information is known, it should be send to all the relevant people, not depending whether these people are busy.

Now that the aspects of crisis management systems are known it is interesting to see where mobile technology can be applied to help solving these problems. What follows is an overview of examples where mobile phones are used in crisis management systems.

A first example of the use of mobile phones is given in (Datcu & Rothkrantz, 2008). Here the usage of a touch screen in crisis management has been proposed as a way to communicate with people that are eyewitnesses of some sort of accident. When a crisis situation occurs, the system can contact people that are in that area and let them provide extra information that can be used by the different emergency response teams. As a tool to help the user describe the situation they can select different images by touching them on the screen and link them to their location. Different images represent different situations, for instance: an image of a flame means there is a fire at the location. This system can be seen as one that addresses the lack of common ground, because the information is created in a standardised format.

In another system for crisis management where usage of mobile devices is proposed the GPS and 802.11 functionality of mobile devices is used to improve the situational awareness for first responders (Betts, R. W. Mah, Mundo, McIntosh, & Jorgensen, 2005). With this system the responders who are called to the scene of a crisis are being tracked by an observation team via their mobile phones which are equipped with either GPS or an 802.11 network. The responders can also communicate with each other by sharing digital images of the scene and text reports. When a responder collects new information, like a new picture of the scene together with some written comments, the responder can make a new *Report* which is the main unit of data in this system. These reports are then uploaded to a central server in which the location of the responder at the time of uploading is also recorded. All the uploaded reports can then be seen in a virtual environment of the scene of the crisis and can be accessed by the other responders and by the observation team. What is important to note is that Betts *et al.* mention the risk of information overload. It is of high importance to make sure that the right information is delivered at the right people so that responders are not disturbed with unimportant information and can concentrate on their tasks. This research addresses the ineffectiveness of other communication technologies and the issue of the breakdown of information flow. The first because this system can be used with the mobile phone, a device that everyone carries around everywhere. The second because information is stored on a server and people can just download this. This also makes sure that there is a common ground for the people working on the crisis situation.

The overview of this research shows the problems faced in crisis management and gives some examples of mobile applications that can be used to address these problems. What problems will be addressed in this thesis will be stated in section 3. The next section will describe the problems with human olfaction.

2.2 Human olfaction

Since the system discussed in this thesis will use people as smell sensors, it is important to know what challenges there are with odor perception. In the excellent master thesis of Joseph Kaye (Kaye, 2001) a good overview is given on human olfaction. The main challenges indicated in this thesis on the human ability to classify odors are *a*) "our system of identifying smells, which involves pointing to objects in the real world and saying "This smells like that"". So people have a strong connections between a certain smell and certain objects in the real world. *b*) Kaye discusses significant evidence that there can be a difference in what two people smell when they are exposed to the same smell. So what you smell can differ from what your neighbour might smell. *c*) Classification changes with experience, as is the case with wine experts who have a better recognition for certain wine smells. *d*) Classification depends on the cultural background of people. As an example Kaye gives the classification scheme of the Serer Ndut of Senegal, who have five different categories including "urinous" which contains Europeans, monkeys, squash and leaves. Although this may be an extreme example it indicates the problem which would occur when a European and a Serer Ndut would have to describe a urinous smell. It seems fairly unlikely that Europeans would describe this smell as "European", where for the Serer Ndut this would be very normal. *e*) Research has shown that women are better in classifying odors, especially pregnant women. *f*) There is the problem of adaptation to certain odors. When a person is exposed to a certain smell very frequently, the person will experience the smell as being much weaker then when experienced before getting used to it. *g*) A connection exists between odor and memory where a certain odor can evoke a, sometimes very old, memory. For example, people can improve their memory when the same odor is present at time of learning as well as at time of recall. Another interesting issue discussed in

Kaye's paper is the phenomenon called *the power of suggestion*. Here Kaye refers to a research that did an experiment where people were informed that a odor, which was described very broadly, was emitted over the television and radio, and that they should call if they smelled something unusual. This resulted in 179 responses after the television broadcast and another 49 responses after the radio. This implies that people can smell something unusual if they are told that there is something to smell, when in reality there is nothing.

In (Wilson & Stevenson, 2003) a case is made for the central role of memory in olfactory perception. They say that "the analytical processing of odors is inaccessible at the behavioral level and that all odors are initially encoded as 'objects' in the piriform cortex". They continue to say that odor perception is completely based in this memory system and that if it were lost it would greatly impair olfactory perception. Some research is described in which people were asked to identify the different components of a certain odor. They had been given the labels for all the odors of which the composition would consist, so in principle had all the tools for giving the components of the odor. Their results show that people find it extremely hard, if not impossible to perform such a task successfully. In other research they find that when a familiar and unfamiliar odor are mixed, the characteristics of both are also mixed. For example, a cherry smell combined with a smoky smell would later result in the cherry smelling more smoky and the smoky more cherry-like. Thus, a connection between the two odors is learnt and properties of the odors interchange. Finally, experiments with amnesiacs show that people without the use of their memory can not differentiate between different odors: for them all odors smell alike. From these researches they conclude that "memory plays a fundamental role in odor discrimination and perception".

Another example of the role of memory in human olfactory perception is given in (Brewster, McGookin, & Miller, 2006) where an application for searching in digital photo collections is discussed. The idea is that the photo collections of people get bigger and bigger which makes searching for specific photo's more difficult. Since olfaction is linked with memory, searching for photo's based on a certain odor would make it easier to find a specific photo where the odor is the tag for the memory of the event displayed in the photo. Next to this, people would learn the connection between a certain odor and a photo in the tagging stage, where the odor is assigned to a photo. This could also help in finding this photo later. While developing this system, Brewster *et al.* found similar problems with odor classification as the ones mentioned by Kaye. When selecting different odors that could function as photo tags, subjects were asked to classify different odors so the experimenters could use the ones which were classified the same by the most subjects. This experiment showed that for the same odor, a lot of different classifications were given. This underlines the difficulties in creating a classification scheme for odors. The final evaluation of this particular application shows that people, in general, experienced the task of labelling and retrieving photo's with the use of odor as a bigger workload. Still, the participants were able to carry out the task successfully, although it required more time and effort. This shows that, given the right odors, they can be used for such tasks because people are able to distinguish them well enough.

A final example of the link between memory and olfaction can be found in (Stockhorst & Pietrowsky, 2004, p.6) where olfaction is described from a more physiological and behavioral viewpoint. A good explanation on why the perception of odors is so influenced by learning is given there:

> *How we perceive and process odors is much influenced by learning (and thus experience), and learning is regarded as a means to deal with the unpredictability of the chemical world. By learning, an organism is able to associate a certain situation with a certain smell. This might be one reason why humans often label smells by referring to the situation in which they encountered them.*

So the link between memory and olfaction can also be explained from the physiological and behavioral viewpoint and might have something to do with our evolution. A second important learning phenomenon of odors is that organisms can learn a robust aversion towards a certain odor which also makes the organism highly receptive for this odor. For example, when people are told a certain odor might be a health risk they perceive this odor as more intense after being told this information. In contrast with this, people can no longer perceive a certain odor if they were told it was good for their health.

From this overview can be concluded that there odor classification is a hard task for people to do. The names people tend to assign to odors are memory and experience based, so they are different for all people. Because of this, up until now a general classification scheme has not been developed. People are also very sensitive for information they have about a certain smell which can influence the intensity with which it is being perceived. All these problems have to be taken into account when creating this dialogue agent, because the goal is to make a start with classifying an odor. What specific problems have to be overcome follows in section 3.

3 Problem statement

As stated in the introduction, the main goal is to develop an Intelligent Dialogue Agent on a mobile phone that can interview people about what they smell. The agent asks people questions about what they smell and then uses this information to detect anomalies and start the classification of the anomaly. To make the time spent with the agent as pleasant as possible, one of the goals of the agent is to make it user friendly. This is done by making sure that the interview takes as short as possible, so people will not get irritated by the amount of time spent with the agent. Next to this the method of answering questions is made easy by just using ”yes/no” questions which people just have to select by using the touch screen. In this way, people don’t have to use the keyboard to type in responses which saves them a lot of hassle.

The second goal of the agent is to of help in crisis situations. From the literature follows that one of the major problems in crisis management is information mismanagement. More often then not it takes too much time to communicate the right information to the right people. This agent should thus be able to communicate the collected information to the right people in the correct format.

Secondly, it seems that collecting the right information from the people is a challenge as well. When evaluating transcripts from DCMR where people called in to report an anomaly, difficulties with this way of extracting information came to light. One of the difficulties is that people tend to elaborate a lot when calling to complain. People probably feel a little irritated by the fact that they are disturbed by a stinky smell and they want to complain about this as well. Some examples of this can be found in table 1 [2]. So people who call do this not only to inform about the smell, but are also telling the operators what they were doing before the were bothered by the smell. They probably also feel that they have taken some time out of their day to help and they feel might experience this as an extra effort. This is all understandable, but is not a very efficient way to collect the relevant information. By simplifying the interview and by limiting it to just ”yes/no” question, people might feel that helping is less of an effort. It also stops people from going into details that are not relevant for the process of gathering information and thus gets the right information in a fast way. The third goal for this agent is to have a good way to communicate with people about odors and their olfaction. From the literature follows that there is no standard classification scheme for odors, which is a big problem for this particular agent. Because the classification for a certain odor can differ from person to person, the goal for this agent is to develop a classification scheme that is as general as possible so that the largest amount of people will classify a certain odor the same way. A solution for this problem is to use questions that evoke memories in people by referring to situations that a lot of people will have experienced in their lives. This follows from all the literature on olfactory perception where the link between odor classification and memory is mentioned. This format of asking questions thus seems a good way to have a better chance of people responding the same. Since the memories that are being tried to evoke should be as general as possible so that a lot of people have experienced them. The format of questions will have to be something like ”Does this smell remind you of. . . ”.

The process of classifying a smell should result in a very general classification where the general category of smell is established. These categories consist of general descriptions of types of gases, an example is ”Chemical”. When these categories are known, the agent has a smaller number of hypotheses and thus less candidate gases, namely the gases which belong to these categories. Because the best

[2] The original transcript is in Dutch, what follows is the best effort of the author to translate this to English

M: operator of DCMR, B: Person calling to complain
[. . .]
M: How long has the stench been bothering you?
B: Well, I'm in the living room and I think by myself what do I smell?
M: Oh, in the living room, so it really got your attention it came from...
B: Yeah.
M: ...outside.
B: Yes, because fifteen minutes ago I opened the bedroom windows and then I didn't smell anything.
M: Yeah.
B: Then I'm in the living room and I think, what do I smell? You know ... gas? No... No... Then I just opened the window
[. . .]
[. . .]
M: Good day.
B: I was just biking in Westwijk.
N: Yes.
At least, I just stepped outside at work, because I was planning to bike home and it really smells there. And I work at child day care so we want to know if the kids can go outside safely.
[. . .]

Table 1: Two examples of unnecessary elaboration

possible estimate on what gas is released should be the final output of the agent, determining these categories can not be the end of the questionnaire. The agent should continue to ask more specific questions about what people smell. Here the problem of odor classification reemerges because the more specific the agent becomes in the questions about what people smell, the higher the chance that people will perceive the odor differently. This means that the more specific the questions get, the more difficult classification becomes and results will become less accurate.

The last problem with the development of questions that follows from the literature is that of the "power of suggestion". Some examples of this can be found in the transcripts of calls to DCMR (table 2. Here two examples can be seen where the operator might "put words in the mouth" of the people who call. This is the result of the knowledge the operators get about the crisis situation and they know what kind of information they can expect. For example, when they have been called about a rotten egg smell ten times already, they presume that the next caller will report this as well. Although this is an very logical adaptation by the operators, it might make people smell something they don't actually smell.

M: operator of DCMR, B: Person calling to complain
[. . .]
M: Erm, let me see, what kind of stench do you smell? A chemical smell?
B: Yes, some sort of chemical smell indeed.
M: And does it remind you of something?
B: Yes, sort of a, yeah everything always smells like paint to me, but well, that is probably me
[. . .]
[. . .]
M: and your description of the stench?
B: Well a terrible chemical smell.
M: A chemical smell? Not rubber?
B: Yeah, a little rubber-like I guess... Yeah...
[. . .]

Table 2: Examples of "the power of suggestion"

As can be seen in the first example in table 2, the operator first suggests the option that the smell might be chemical. Hereby the operator might put this idea in the mind of the person who is calling, while this person actually smells something different. The second example in table 2 is an even clearer case of this phenomenon. Here the caller tells the operator he smells something chemical, after which the operator immediately suggests a rubber-like odor. Again, this might make the caller think he smells rubber, since it was suggested, all the while he might smell something different.

Summarising this section, the following requirements for the Intelligent Mobile Dialogue Agent can be distinguished:

- The output should be information that can be used immediately by the appropriate experts.
- The dialogue should be guided and simplified to avoid unnecessary elaboration and get the relevant information as fast as possible. This will:
  - save time
  - make the process as efficient as possible
- When talking about olfactory perceptions:
  - avoid suggestive questioning as much as possible
  - refer to the memory of the user

In the following section the solutions to these problems together with their implementation will be given.

4 Implementation

This section will discuss the solutions to the problems stated in section 3 together with their implementation in the Mobile Intelligent Dialogue Agent. Every problem will be discussed separately.

4.1 Interface

One of the goals for the agent is to make the user's experience while working with the agent as pleasant as possible. Therefore, some consideration in the design of the interface is important. The interface is implemented in Android [3], the new operating system for mobile phones by Google. This operating system runs on phones with touch screens, which means that the previously mentioned option to use this technology is very easy.

Another design choice is made with respect to the dialogue about olfactory perception (see section 4.2). Because the dialogue should be kept as simple as possible, there will be only questions asked which can be answered with "yes" or "no". The user will see a question followed by two buttons labeled yes and no which can be selected. After the answer is chosen, the user just has to press the "next" button to see the next question, which is selected with respect to the answer given (see section 4.3). An example of the layout can be seen in figure 1. In this way, the interview is guided and kept simple so the user has no chance to elaborate. Instead, the user has to focus al the attention to answering the questions asked by the agent, which should result in faster questioning and higher quality of the retrieved information. When people still feel the need to elaborate on what they smell or have anything else to tell to the agent, an opportunity for this is given as the last question. Here people are given an empty text box in which they can type whatever they want to say, and this information will then be stored with the possibility to be evaluated by operators or other experts later. People will be told that they have this opportunity in the introduction of the agent.

[3] http://developer.android.com/index.html

Figure 1: Example of the layout for a question

4.2 Olfactory perception

The classification of odors by people is the second problem for which a solution has been found. To have an effective way to ask people about what they smell, a lot of attention is given to the questions that the agent should and should not ask. This is done by starting to categorise different kind of odors into more general categories. By doing this, people are first asked very broad questions about what they smell, while this still helps in narrowing down the number of candidate gases in the hypotheses space. The categories are extracted from a document of DCMR in which stench classifications are given, see table 3. DCMR has a document in which general categories are made for different kinds of odors. This document is divided into big categories such as "Chemical smell" or "Sharp smell" and within these categories some more specific categories of the big category are given. For instance, some of the more specific categories within "Chemical smell" are "ammonia", "rubber" and "plastic". These categories will be used for general question selection, and for classification (see section 4.3.1). When

General categories	
Anomaly	Burning
Chemical	Rotten
Disinfectant	Heavy
Oil(y)	Sweet
Mercaptan	Sharp
Hospital	Gas station

Table 3: General categories from DCMR information

asking questions about these categories, links with the memory of the user are tried to be made as much as possible. For example, when the agent wants to know if the smell is in the "mercaptan" category (natural gas is part of this category), the question is: "Does the smell remind you of your stove?". This question tries to evoke the memory of that one time the user waited a bit too long with lighting the gas and smelled the typical gas smell. By doing this, it should be easier for the user to classify the current odor as being similar to the odor linked to the memory, or not at all. This way of starting the conversation about olfactory perception also reduces the possibility of the power of suggestion, since the first question are so general that people have to make up their own mind on what they smell. How the selection of questions for this phase of the interview is done can be found in section 4.3.1

After the first questions are asked, the general category of the smell that people perceive is determined. Then the hypotheses space of candidate gases has been reduced and the next goal is to get the hypothesis with the highest probability. This can then only be done by asking more specific questions about the odor and here the power of suggestion is hard to avoid. To be able to ask more specific questions about specific gases, some sort of database is needed where gases are described with: *a*) their molecular name *b*) their name in every day life *c*) as much specific smell properties as possible *d*) (if present) as much physical complaints as can emerge when exposed to the gas *e*) the general categories in which they belong At the moment of writing this thesis, such a database does not exist. Therefore, one has been created using both the previously mentioned information from DCMR and the world wide web. Some example gases have been selected, after which a search for as much smell properties and physical complaint properties was done on the internet. This has resulted in a small database with fifteen gases and their properties (see table 5 in the Appendix). Further work on such as database seems evident to be an important subject for future work (see section 7).

The gases are then represented as a Bayesian Believe Network, where each gas is a node of which the entropy should be made as low as possible. An exact explanation on how the questions are selected to reduce entropy the most is given in section 4.3. Examples of questions in this phase of the conversation are: ”Does it smell like strawberry?” and ”Does the smell remind you of perfume?” or ”Is your skin irritated since the smell started?” and ”Does the smell give you a headache?”.

4.3 Question selection

This section describes the intelligent component of the Mobile Intelligent Dialogue Agent. With this component the agent is able to select questions that in the first phase of the interview narrow down the current smell into some general category as fast as possible. In the second phase questions are selected that reduce entropy for the gases that remain after the first phase, so the ones that belong to this category.

4.3.1 General questions

As stated in section 4.2 there are some general descriptions for describing categories of odors, see table 3. These descriptions have been used to make a Bayesian network where highest nodes are the most general and the lower nodes get more specific (see figure 2). As can be seen in the figure, the top most node is labeled ”Anomaly”. This is the first and most important piece of information when starting an interview, the question that is asked can be seen in figure 1. When the user does not smell anything unusual, the interview does not have to continue, while very useful information has been collected. The DIADEM system is primarily interested in how many people smell an anomaly. When this number is bigger then a certain threshold, an investigation with chemical experts is started. The rest of the interview for classifying this anomaly is also useful, but the anomaly question is the most important. After this first question is asked, the second question is selected from the next four possible categories. This is done by checking which category is the most likely, given the value of it’s parent (in this case the ”Anomaly” node). Each node in the network has a probability table which is initiated with values based on historical data. Large amounts of data are used that have been collected from previous chemical accidents. This data is labeled with the gas that was actually released at that time, and it contains all the information people provided DCMR at the time of that incident. This data can be used to learn the probability tables of the nodes, based on the number of times people answered questions a certain way. For example, when the gas chlorine was released, people told the operators at DCMR that they smelled a swimming pool-like smell 95% of the time. These quantities are used to learn the values of the probability tables by using for instance gradient ascent (see equation 1). Learning these probability tables is a topic of current research, see section 7. For the purpose of this research the probability table is based on intuitive values where common sense was used to ”guess” values for these tables, for an example see table 4. This table says that the chance that the odor is a member of the category chemical is 0.7 given that there is an anomaly. These numbers are based on the fact that most of the time the

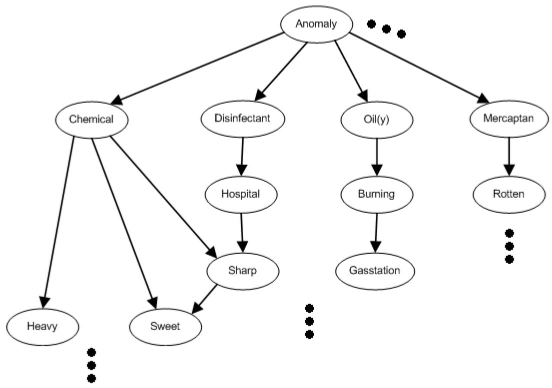

Figure 2: Visual representation of the partial Bayesian network for the general categories, used for this application

crisis situation will involve a chemical smell, but these values are not based on empirical or historical data. The algorithm is designed in such a way that it can adapt to the current situation by changing these probabilities based on answers of users and in this way ask the most probable question given the current situation. How this is done will be explained later. The agent continues asking question until some external node in the network is reached. Now the agent knows in which of these general categories the odor smelled by the user belongs, so it can select the candidate gases from the hypotheses space which are in these categories. How selection of questions in the second phase is done can be read in section 4.3.2.

All the answers from a user are stored, and after the interview is done this data is used to adjust the weights in the network to fit the current situation. This is done by gathering a set of answers from as many users as possible and using these to update the network using gradient ascent. It should be noted that these changes are not permanent and are only used for adjusting the agent to the current situation, since the learning of the actual probabilities requires a lot of labeled data as mentioned before. This means that the data should contain the name of the actual gas that was released when the information was gathered. The adjusting of the weights done here is for experimental purposes to see if the agent can adapt to the current crisis situation, so it can ask the more probable questions before the less probable. Gradient ascent training has been implemented as described in (Mitchell, 1997). This way of training

Chemical		
	Anomaly	$\neg$ Anomaly
Chemical	0.7	0.05
$\neg$ Chemical	0.3	0.95

Table 4: Probability table for the node "Chemical"

maximises $P(D|h)$, which is the chance of hypothesis h occurring given some set of training examples D. By maximising this term, the entries in the probability tables are adjusted with respect to their parents, which define the probabilities for these entries. If w_{ijk} is an entry in a certain probability table, its value is defined by the value of its parents U_i. The term w_{ijk} denotes the probability that node Y_i takes the value y_{ij} given that the parents U_i take on the value u_{ik}. Translated to this problem this means that the chance that the node "Heavy" takes on the value "true" is defined by the value taken on by its immediate parent "Chemical". The update rule used to adjust the weights in this network therefore is:

$$w_{ijk} = w_{ijk} + \eta \sum_{d \epsilon D} \frac{P_h(y_{ij}, u_{ik}|d)}{w_{ijk}} \tag{1}$$

Following the previous example, $P_h(y_{ij}, u_{ik}|d)$ is the number of examples d form D where "Chemical" is true and whatever value y_{ij} then takes. The parameter η is the learning rate, a number between 0 and 1 which determines how much influence each update has. Finally, the resulting probability has to be normalised, so the summed values of y_{ij} equal 1. By updating all the probability tables, the network should adapt to the answers given by users. So for example, initially the chance of the smell being "Chemical" after an anomaly is the biggest, but a lot of users say the odor smells more like a disinfectant. This should be processed by the agent by learning from these answers and in this way adjust the tables to make the probability of "Mercaptan" higher and the probability of "Chemical" lower. Results of this implementation can be found in section 5.

4.3.2 Specific questions

Firstly, it is important to note that the functionality described here has not yet been implemented in the current version of the agent, this is left as future work (see section 7). What follows is the theoretical idea behind this method, together with a simple example with only two gases implemented in Hugin Lite, a tool for creating Bayesian Networks and the appropriate algorithms for learning and inference [4].

The final output of the agent that can be used by the experts and decision makers should be the previously mentioned presence or absence of an anomaly and the probability for the hypothesised gases. After the general questions have determined which gases can still cause the smell, the next goal is to reduce the amount of uncertainty of the question if a specific gas is present or not. This is done by selecting the questions that reduce the *entropy* of this uncertainty. A Bayesian Belief Network is again used to implement this procedure, figure 3 shows the example network created to illustrate the idea. Since

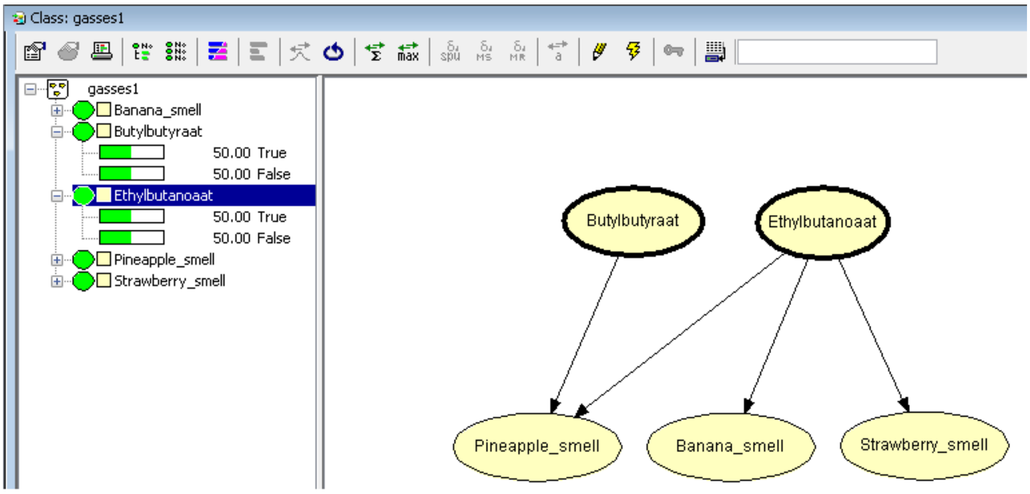

Figure 3: Representation of the BBN with two example gases and their corresponding properties. This network was created with Hugin Lite. (http://www.hugin.com/)

this is the same kind of network as used in the general category classification, all the nodes have a corresponding probability table. For example, the node "Pineapple smell" contains the probabilities that this smell is present given that "Butylbutyraat" or "Ethylbutanoaat" or both are present. Again, the problem with initializing these probability tables is that they have to be learnt from large bodies of labeled data, so again the values for the probabilities are based on common sense for the purposes of this project.

Now, to select the questions that reduce entropy the most, a simulation must be done to see *what if* certain odors are present or not. In the case of the example, an algorithm should see what would happen if the three different gases are being perceived. As can be seen in figure 3, the probabilities of the two gases being present or not is 50/50. This means that the entropy in this case is 1, because the chance that

[4] http://www.hugin.com/

a gas is present or not is equal to the flip of a coin. By initializing one of the property nodes with ”yes” the gas is present or ”no” it is not, inference within the network can determine what this would mean for the chances of a gas being present. In figure 4 can be seen what the difference in probability would be if the value for ”Pineapple smell” would be false. As can be seen, this bumps the chances of either of the gases being not present up to about 71%. This thus has reduced entropy to 0.4 ($\frac{28.57}{71.43}$), which means a Δentropy of 0.6. Repeating this procedure for all possible questions results in a node that can reduce entropy the most and increase Δentropy the most. This node then represents the question that should be asked by the agent to gather as much information about what gas is being smelled by the user. The interview would then stop when some threshold of entropy is reached, which should be determined by the experts that will use this information.

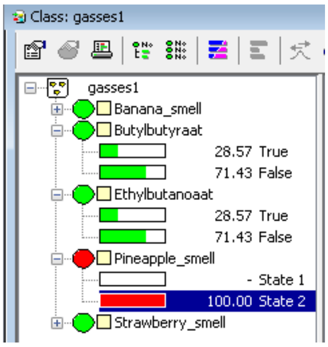

Figure 4: Situation of the network when the smell of pineapple would not be present

5 Results

This section will present all the results of the implementations that have been done. Also, some small tests with the Bayesian network will be discussed.

5.1 The Mobile Intelligent Dialogue Agent

The main result of this project is the first implementation of a Mobile Intelligent Dialogue Agent that can interview people in the case of a crisis situation. The Agent can be used in case of an industrial accident where a gas is released. The Agent is able to ask people question about what they smell in such a crisis situation. To illustrate this, an example of a scenario where the Agent could be used is given in the Appendix (section 10).

As can be seen in this scenario, after only one call has been made to the control room of DCMR, the agent can be activated to contact all the people near the area of the first complaint who have the Agent installed on their mobile phone. After people are contacted, the Dialogue Agent starts an interview to see if the people who where contacted smell some anomaly. The Agent is implemented in Android and some screen shots of the application can be seen in figure 1. The Agent selects questions based on a Bayesian Network where every node represents a category in which the smell can belong. These categories are based on information from DCMR and they are given in table 3. The probabilities of these categories being true given the value of the parents of the category are represented in probability tables. The values in these tables are used to select the most probable next questions, given the answers of a user. In the current version of the Agent these tables are not based on any historical empirical data, but are initiated according to common sense. This is good enough for the purpose of this research, since all the functionality of the Bayesian Network can be tested with these values.

The Agent continues asking questions until a leaf node (one of the nodes without children) in the network is reached. When this is the case, the smell is classified into this leaf node category, plus all the categories where the user answered ”yes”. For example, when the interview stops at the node ”Rotten”, the smell is classified as belonging to the categories of ”Anomaly”, ”Mercaptan” and ”Rotten”. In this way, the Agent has reduced the number of possible gases to those that belong to all these three categories and thus a start of the classification of the smell is made.

When a leaf node is reached, some more specific question about the category of the leaf node are asked. These are questions that must specify the properties of the gas. For example, when the leaf node that was reached in a certain interview was ”Sweet”, questions like ”do you smell pineapple?” are being asked. In the current version of the Agent all questions belonging to this category are asked. As discussed in section 4.3.2 this can be extended to use Bayesian inference, see also section 7.3. With the answers of these specific questions the Agent can then continue to classify the smell by checking which gases possess these properties. The data that is generated by a user is then saved and used to adapt the Bayesian Network to the current situation. This is done by accumulating the answers of for example one hundred participants and giving this data to the network. The results of this learning will be described next.

5.2 Bayesian Network

As stated in the previous section, the agent can learn from user generated data to adapt the probability tables to the current situation. By this is meant that when the initial guess of a certain node does not apply in the current situation, because people smell something else than is expected, the probability table where wrong. Then the Bayesian Network should adjust the values so other nodes be more probable. For example, first the ”Chemical” node has the highest probability to be chosen after ”Anomaly” with 70% chance. But 90% of the people say they don’t smell anything chemical, but the say that they smell something gassy. In this case, the probability of ”Chemical” being true after ”Anomaly” is decreased and the probability of ”Mercaptan” being true after ”Anomaly” is increased.

To test this principle, some simulated should be given to the network to see if it could adapt the probabilities of the nodes correctly. Because of the way the Bayesian Network is implemented in the current version of the Agent, it is hard to feed the network with this data and collect results from these tests. The connection between the Bayesian Network and the application is made using a server/client connection. This makes it hard to feed the network large amounts of generated data, since the answers are collected from the application. So to make one hundred simulated interviews would mean to start the application one hundred times and select the appropriate questions. This connection between the application and the network was chosen because of the opportunities this would give for user experiments. However, it turned out that this way of implementation is not ideal for these tests.

But because of the learning rule implemented in the Bayesian Network is, it is able to learn to change the probability tables of nodes. For test purposes the hypothetical situation from the scenario is simulated where the gas ”ethylbutanoaat” is released. This gas belongs in the ”Chemical” and ”Sweet” categories, so the initial probabilities of the tables asking for these categories are highest. Then data is generated where people do not smell something chemical, but something gassy by letting 80% of the simulated data answer the question ”Would you describe the smell as chemical?” with ”No”. The other 20% does respond with ”Yes” to the chemical question. One hundred simulated interviews are made and then should be fed into the Bayesian Network. The network will then see for every visited node what answer was given and what values the parents of this nodes had at this time. So in this case for the node ”Chemical” eighty interviews of the simulated data answered ”No” to this question while node ”Anomaly” (the parent) was true. The update rule from equation 1 is then applied which increases the value for the entry that says ”Chemical” is false given ”Anomaly” is true. Then the values of the table are normalised, since the should add up to one, which results in a lower value for ”Chemical” being true after ”Anomaly”. In the same way, the value for ”Mercaptan” being true after ”Anomaly” is increased. In this was the Agent will give the questions of ”Mercaptan” before that of ”Chemical”. However, no empirical evidence for this can be given here since feeding the data to the network was difficult because of the implementation.

The Bayesian Network can however choose the most probable next questions by looking at the probabilities of the different next possible nodes. All the different paths in the network have been tried with different values for the probability tables and every time the correct question was generated based on these values.

6 Discussion

How good is this implementation of the current version the Mobile Intelligent Dialogue Agent? As can be seen in the results most of the desired functionality described in section 3 has been implemented in the current version. Some of the goals of the agent have not yet been tested. For instance, to see if the Agent actually is faster in gathering information it should be tested versus people. It does seem however, that a guided and simplified interview like the one done by the Agent will be faster than a human interview over the phone. Because the system is pro active, it contacts people instead of waiting for people to contact him, more people can be interviewed within less time. Another goal of the agent, the ability to ask to most probable questions based on probability tables first, is not tested completely yet either. But resigning a second version of the Agent would make it fairly easy to feed the network with the generated interviews and see how good the Agent could adapt to a crisis situation.

A second big part of this project is to investigate how people perceive odors and how they use their olfactory system to classify these odors. Here a lot of interesting facts have been discovered that can be very useful for improving the information gathered by the Agent. The use of general categories is a good way to start classification of an odor since these categories are based on information from actual crisis situations and descriptions people gave there. The link between memory and olfaction is also very important, because the classification a person gives to a certain odor is often based on a memory or on a certain experience. This information is implemented in the Agent by asking questions that refer to memories people might have had when they smelled a certain odor. For instance, when asking if people smell natural gas, the Agent refers to memory by phrasing the question as ”Does the smell remind you of your stove?”. This can evoke a memory in people’s heads of occasions in which they smelled the natural gas of their stove which makes classification for this smell more accurate.

The limitations of the Agent together with suggestions for future work will be discussed in the next section, after which a conclusion of the implementation of the Agent will be given.

7 Limitations & Future work

As previously mentioned in sections 4 and 5, there are some limitations to the implementation of this version of the Mobile Intelligent Dialogue Agent. This section will discuss these limitations and describe suggestions for future work.

7.1 Gas database

The first limitation discussed is the gas database that is used in the current version (see table 5 in the Appendix). This database now consists of only fifteen gases and the properties that are now listed are from some websites on these gases. These are not the most reliable sources for this kind of information, and so the current list is probably not ready for use in a real world application. To improve this list, cooperation with organisations like DCMR seems very relevant. These kind of organisations probably have the expertise and knowledge in this field to extend this database. The document of the smell descriptions is one example of resources that DCMR possess, so for them an extension for this document in the form of a gas database must be relevant. During the course of this project DCMR has been contacted but as of this moment no progress has been made on the database. Future work on a next version of the dialogue agent would include extending this database in close cooperation with DCMR.

7.2 Probability tables

The second limitation is related to the first in the way that data from cooperations such as DCMR is needed to be able to implement this feature. This concerns the learning of the probability tables for the nodes in both of the used Bayesian Networks. The original idea was that this agent should be able to ”learn from experience” by using the results from the interviews to learn the probability tables. This function has been partially implemented in the current version, but this does not cause permanent changes in the probability tables and is just for a current crisis situation. As a result of this, the probability tables used in the current version are, although based on common sense, made up. In future research it would be good to collect as much labeled historical data from other crisis events and use this to learn the ”actual” probability tables. Attempts to do this have been made in (Schuppen, 2008) where some hopeful results are reported. But since then more data has been collected and it seems worthwhile to make another attempt to improve this work. If successful, this would make the application more suitable for real life deployment.

7.3 Entropy reduction

As was previously mentioned in section 4.3.2 the entropy reduction algorithm has not been implemented in the current version. This is because first an attempt was made to learn the probability tables from the interviews. This turned out to be both impossible and undesirable. Impossible, because this can only be achieved with large amounts of labeled data, as stated in the previous section. Undesirable, because the goal of this agent is to classify the gas for which the property tables of the nodes in the network must be know and can not be changed. The entropy reduction function is now only described in theory, but for future work this element should definitely implemented.

7.4 Confidence interval

The last function that might be interesting to implement in a future version of the agent is the addition of a confidence interval for users to express their confidence in their answers. This functionality seemed relevant while developing the network for the general categories and the question came up what to do when people where still unsure of what they smelled. In this case, it would be nice if the user has the possibility to express how sure she if of her answer. This indication could then be used by the question selection algorithm, for example to decide that some more questions should be asked to the user to lower the uncertainty.

7.5 User tests

The last limitation of this research is that no user tests have been done as of this moment. One possible user test is to let people select questions for users and let the dialogue agent do this. This would show if the agent is actually better in selecting questions then a person and if the agent can do the interview faster and make better results. Other tests could include testing the system in a real life setting where a certain odor is released and the agent should figure out what odor was released as fast as possible. The application is completely setup to be used for such tests since it contains the previously mentioned server/client connection.

8 Conclusions

As a conclusion can be stated that this first version of the Mobile Intelligent Dialogue Agent is a good start from which better version can be created. The implementation of the Bayesian Network can be reconsidered to make testing with large amount of (generated) data easier. The current Bayesian Network can choose the most probable questions based on the answers of people by using the probability tables of the categories. In the field of human olfaction very interesting facts have been discovered that will help in classifying gas better and faster within a crisis situation.

9 Acknowledgements

- Vanessa Evers, for being an excellent supervisor and always comming up with new idea's.
- Andi Winterboer, for being an excellent supervisor as well and always being in the audience while I presented my research.
- Mattijs Ghijssen, for helping me a big way with coding in Android by pointing out beginners mistakes.
- Zora Jurriëns, for being the hostess of several late writing sessions and for being such a good motivator.

10 Appendix

10.1 Scenario

This scenario contains a description of a situation in which an intelligent dialogue agent can be used to extract information from participants. The situation is that a certain factory has some gas leak of which they are unaware. Participants are contacted over their mobile phone and asked some questions about what they smell. This information can then be used to determine, with a certain level of confidence, what gas is released:

Step 1:
A factory 'A' that produces chemicals for industrial use is situated near a residential area. Around this area are a number of other factories. At some Friday at 3 PM, the seal of one of the containers in factory 'A' that holds a gas used for production starts to leak. The leak is not detected by the factory staff, so for some time the gas spreads towards the residential area. The gas is Ethylbutanoaat, one of the properties of the gas is that is has a very chemical and sweet smell, a little bit like perfume or pineapple.

Step 2:
A woman who lives in the residential area whose house is very close to factory 'A' is hanging her laundry in here garden at 3:05 PM. She suddenly smells a weird smell, which get's stronger by the minute. She decides to call the "Meldkamer" of DCMR, who she knows should be contacted if she smells something strange.

Step 3:
At the "Meldkamer" of DCMR the call from the woman is received and here complaint is recorded. There have been some other similar complaints the last couple of minutes, so DCMR decides take action and see if something is wrong. A large group of volunteers living in the same area as the woman who called, is contacted via their mobile phone where the "Mobile Intelligent Dialogue Agent" application is activated. This agent will interview the volunteers about what they smell and use their answers to generate new questions. The agent has an intelligent component and access to a database of gasses with some of their properties. This component can select questions about properties of the possible gases that can discriminate between the different gasses as fast as possible.

Step 4:
The mobile phone of one of the volunteers rings. The owner picks up and see's that it is the agent to ask him about what he smells. First he is informed about the current situation and that his help would be very valuable in determining what gas is being released. Then he is presented with the information about his location which is known by the agent. The participant is asked to check this information and correct it if it is wrong. Then the man is asked some questions about what he smells. This are questions like "Do you smell anything unusual?", "Does this smell remind you of perfume?" and "Does the smell irritate your eye's". He can answer these questions with "yes" and "no" and his answers are used to generate new questions. When the interview is done, the information is being sent back to DCMR, who now have his answers and the resulting gas that the intelligent component has figured out.

Step 5:
When all the information of all the different volunteers is collected at DCMR. They can use this information in the rest of the process of determining the source and kind of gas that was released and all further steps.

10.2 Gas table

Gas	Spreektaal	Eigenschappen	Categorie
Ammoniak	Schoonmaak middel (vooral glas)	Penetrante lucht Irritatie ogen	Chemisch Scherp
Chloor gas	Zwembad lucht	Zware chloor lucht Irritatie luchtwegen Irritatie ogen Metaal smaak	Chemisch Zwaar Scherp
Butylbutyraat	Snoepjes Fruitig	Ananas geur	Chemisch Zoet
Ethylbutanoaat	Snoepjes Zeep Parfum	Ananasgeur banaan geur Aardbei geur	Chemisch Zoet
Ethylheptanoaat	Fruitig	Druif geur Kers geur Abricoos geur	Chemisch Zoet
Isobutyl acetaat	Fruitig	Kers geur Aarbei geur Misselijkheid Hoofdpijn	Chemisch Zoet
Methylantranilaat	Fruitig Parfum	Druif geur Jasmijn geur	Chemisch Zoet
Methaantiol	Slechte adem/scheet	Rotte kool scheet slechte adem	Mercapatan Rot
Aardgas	Geur van het gasfornuis "Gewoon gas"	Hoofdpijn Typische gaslucht	Mercapatan Rot
Benzine	Auto lucht	Brandstof geur Verbrande olie geur "Benzine geur"	Olie Brand Tankstation
Diesel	Auto lucht Scheepvaarts lucht	Brandstof geur Verbrande olie geur "Benzine geur"	Olie Brand Tankstation
Etheenoxide	Ziekenhuis ontsmetting	Duizeligheid Hoofdpijn Irritatie huid	Ontsmetting Ziekenhuis Scherp
Dimethylether	Spuitbus lucht	Duizeligheid	Ontsmetting
Diethylether	Ziekenhuis lucht	Ziekenhuis geur	Ontsmetting Ziekehuis
Anisol	Anijzerig	Irritatie huid Irritatie ogen Duizeligheid Anijs geur	Zoet Scherp

Table 5: Gas database created for this project (in Dutch)

References

Betts, B. J., R. W. Mah, R. P., Mundo, R. D., McIntosh, D. M., & Jorgensen, C. (2005). Improving situational awareness for first responders via mobile computing. *NASA and QSS Group Inc*.

Brewster, S. A., McGookin, D. K., & Miller, C. A. (2006). *Olfoto: Designing a smell-based interaction.* (CHI 2006 Proceedings)

Datcu, D., & Rothkrantz, L. J. (2008). *A dialog action manager for automatic crisis management.* Washington DC, USA. (5th International ISCRAM conference)

Kaye, J. N. (2001). *Symbolic olfactory display.* Unpublished master's thesis, Massachusetts Institute of Technology, USA.

Mitchell, T. M. (1997). *Machine learning* (International ed.). The McGraw-Hill Companies, Inc.

Pavlin, G., Wijngaards, N. J. E., & Nieuwenhuis, C. H. N. (2009). *Towards a single information space for environmental management through self- configuration of distributed information processing systems.* Masaryk University. (TOWARDS eENVIRONMENT Opportunities of SEIS and SISE)

Reddy, M. C., Paul, S. A., Abraham, J., McNeese, M., DeFlitch, C., & Yes, J. (2009). Challenges to effective crisis management: Using information and communication technologies to coordinate emergency medical services and emergency department teams. *International Journal of Medical Informatics*(78), [259-269].

Schuppen, D. van. (2008). *An investigation of the dcmr domain for the automated compound detection.* Unpublished master's thesis, Universiteit van Amsterdam, Amsterdam, The Netherlands.

Stockhorst, U., & Pietrowsky, R. (2004). Olfactory perception, communication and the nose-to-brain pathway. *Psychology & Behavior*, [6-7].

Wilson, D. A., & Stevenson, R. J. (2003). The fundamental role of memory in olfactory perception. *TRENDS in Neurosciences*.