Institut National des Sciences Appliquées de Rouen
Laboratoire d'Informatique de traitement de l'Information et des Systèmes
Universitatea "Babeş-Bolyai"
Facultatea de Matematică şi Informatică, Departamentul de Informatică

Extended Abstract for
P H D T H E S I S

Speciality : Computer Science

Defended by

Ovidiu Şerban

Detection and Integration of Affective Feedback into Distributed Interactive Systems

PhD Directors:

Horia F. Pop	-	*Professor*	-	"Babeş-Bolyai" University
Jean-Pierre Pécuchet	-	*Professor*	-	INSA de Rouen

Contents of the thesis

Keywords: *Affective Feedback, Human-Computer Interaction, Emotion Detection, Contextualised Dictionaries, Storytelling Environment, Distributed Interactive Systems*

Introduction

Human-Computer Interaction

The interaction paradigm assumes the continuous reciprocal influence between two individuals. From another perspective, it is a game where one acts whereas the other reacts. In computer engineering, the interaction takes place between a computer and an individual. Historically, this process evolved from a *one way* communication, where the computer was playing the role of executor, to a *bidirectional* communication, with the system as a communication partner. The communication protocol is changing, as well, from the classical master-slave perspective to a collaborative environment.

Traditionally, when the computer is just an executor, the input channels formalised as simple buttons are sufficient to ensure a good level of communication. Nowadays, the interaction becomes more intuitive. The inputs are a fusion of voice, gestures, postures or physical signals (such as acceleration, speed, orientation). The buttons still play a major role in this process, because of their accuracy, but the migration is slowly moving towards naturalistic input.

From the computer perspective, the output changed from the classical text interfaces to graphical and then to more intuitive ones, such as dialogue oriented interfaces. The Ambient Intelligence (AmI) field models scenarios where a computer is able to control the environment in a natural way (i.e. switch on/off lights, execute daily tasks, communicate through spoken language) [52, 173]. Moreover, the computers may have a personified appearance (i.e. an Embodied Conversational Agent (ECA) [155, 40] or a Robot [64]) or even a personality [175]. More formally, we will use throughout this thesis the concept of *agent* to describe an intelligent entity, such as an interactive system.

Current approaches describe the human-computer interaction as a collaborative, dialogue-based, task [6]. Both interaction and dialogue task involve rich exchange of information between at least two peers. The richness of input refers to the possibility to use multi-modal exchange in the communication process: spoken language, gestures, postures, vocalisations. The main difference between the two aspects is that, usually, the dialogue involves knowledge management, whereas the interaction can be strictly

reactive. So far, softer reactive dialogue models have been proposed, without being sufficient to ensure a good communication level [137]. In our work, we refer to *dialogue* as a model where certain level of response planning and knowledge management is involved, and to *interaction* as a model that is mainly reactive.

Both dialogue and interaction models involve feedback detection and generation. Several levels of feedback can occur at any point through the information exchange process [37, 8]. These include: perception feedback (positive if the phrase can be transcribed, negative in case of failure), interpretation (positive if the phrase can be interpreted correctly according to the rules describing the system, negative in case of a misinterpretation) or execution feedback (positive if a satisfactory response is generated, negative otherwise). Out of these, a special category of feedback is represented by the human emotions. They do not act directly at a certain feedback level, previously described, but influence them all. For example, a negative emotion in the context of a perception failure can influence the response style. Instead of replying a simple phrase, such as: *"I am sorry, but I do not understand"*, the agent could also build a solution for the problem: *"I am sorry, but I do not understand. I would increase the volume of the microphone and let you try again."*. In this example, the system is able to detect frustration, as a negative emotion, and propose a solution instead of just giving the result error.

Affect Oriented Modelling

In this perspective, R. W. Picard is one of the first computer scientist working in Affective Computing, to offer a new point of view for engineers [152]. In order to make human-computer interfaces more interactive, she proposed to integrate emotional models into existing approaches. She described the problem not as a strict detection or simulation issue, but with very fuzzy boundaries. One of the challenges underlined by Picard is that such systems would balance the detection rate against the user's satisfaction.

According to Oxford Dictionary [143], the Online Edition, an emotion is a strong feeling deriving from one's circumstances, mood, or relationships with others. On the other hand, the opinions are the beliefs or views of a group or a majority of people. From a general perspective, an emotion is more complex and fuzzy than an opinion. Usually in opinion mining field, the literature refers to the valence (negative or positive) of a certain opinion [28], which is a simplified model of an affective intensity.

From the psychological perspective, P. Ekman proposed his original emotion model [56, 54] using only six basic emotions, considered as universal and recognizable all around the world: Anger, Disgust, Fear, Happiness, Sadness, Surprise. This work is the foundation of the Universality Theory [45], which states that all living beings express emotions in the same way. W. James is one of the two pioneers in the field of physiological perception of an emotion, foundation for most of the signal processing techniques [92].

Even with some controversies in the area, several European inter-disciplinary or-

ganizations decided to launch ambitious projects such as HUMAINE Consortium [86], which aims at linking different research communities, all working on the idea of human centred research. The initial phase of the project finished with the release of the HUMAINE database [49], that contains a video corpus annotated with different schemes, among which a basic set of tools needed to analyse the data. The project continues as an excellence network, with a lot of researchers involved.

More recently, the SEMAINE consortium, as part of the HUMAINE Excellence Network, finished a project that focuses on the multi-modal detection aspect [175]. The Sensitive Artificial Listener (SAL), proposed by SEMAINE, is able to detect human emotions based on face gestures and several qualitative speech features. Besides this, the AVEC Challenges [179] propose a set of annotated corpora to solve the same issue.

Original aspects of this thesis

None of the previous presented works focuses on the semantic level. Humans, in everyday interaction, use natural language, among other modalities, to exchange information. The semantic level corresponds to the information transmitted, to **what** is being exchanged. The gestures, postures and vocal features are linked to the transmission style, or **how** things are being transmitted. We agree that the fusion of multi-modal features is a difficult task, but we also state that the semantic part of the communication has to play an important role into the detection process: the **how** and **what** have to be considered together.

Currently, the semantic context of affective words is very poorly exploited. When it is done, in most cases, manually annotated linguistic resources are proposed. We suggest an approach that deals with context in affective dictionaries which is generated automatically out of linguistic resources freely available over the web.

In order to make the machines "understand" human emotions [152], the algorithms that deal with affect detection and simulation need to be integrated into a system. Moreover, many of the problems regarding affect-oriented interaction systems are currently solved (partially or entirely). Therefore, the integration of all these components into a unified platform becomes critical. Our proposition, AgentSlang, is built around the idea of component integration and provides an architecture for this purpose. Many other important steps have to come, but even so, this platform remains one the biggest contributions of this thesis.

Long Term Goal

The intelligence is characterised by the ability to acquire and apply knowledge and skills [143]. Building an intelligent agent, described by these abilities, is a very difficult task. Nevertheless, such a behaviour can be simulated by integrating feedback detection mechanisms, which would make the whole system more interactive [152].

The final goal of this work is to build a natural interactive environment, by using Embodied Conversational Agents or Robots. The approaches proposed are reactive,

based on the semantic feature extraction and emotions detected in a multi-modal context. The usage of affects would increase the interactiveness of the system, while being able to provide real-time feedback for a dialogue model.

Structure of the thesis

This thesis investigates two main directions: algorithms used to detect user's emotional feedback or to build affective linguistic resources, and systems constructed around the interaction paradigm. Figure 1 presents a detailed structure of this thesis, with links between several sub-projects.

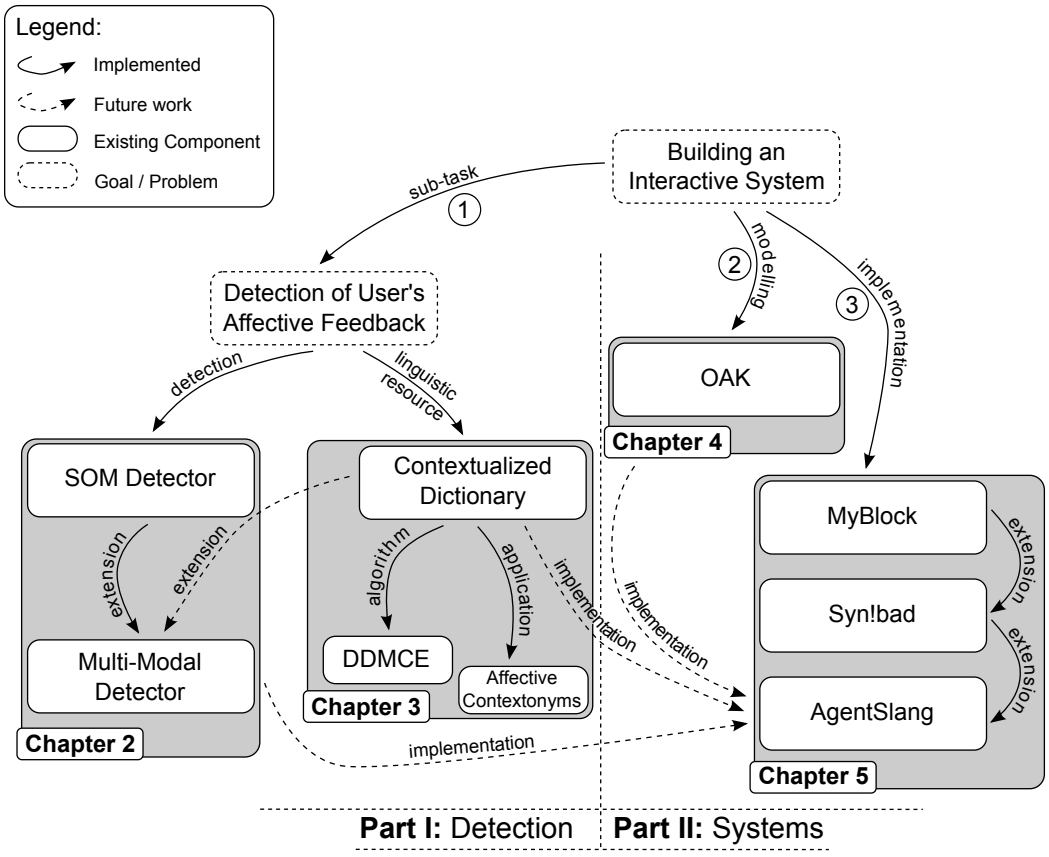

Figure 1: The thesis structure and the interaction between the different parts

As stated before, the goal of this thesis is to build a system that deals with natural interaction. Several aspects of this problem need to be discussed:

(1) The problem of natural interaction is linked to emotions and affect detection. This is one of the major parts of this thesis.

(2) Due to the lack of rich interaction data, especially for children, an experiment dealing with corpus collection was performed.

(3) The design of an interactive system is proposed, which integrates several existing algorithms.

The problem of Affective Feedback Detection is tackled first by a Self Organizing Maps (SOM) Algorithm **(c9)**. This approach led to a second algorithm, which uses multi-modal features and a Support Vector Machine (SVM) to detect emotions **(c3)**. These models are described in **Chapter 2**.

While developing the SOM algorithm, we observed that the existing linguistic resources are not accurate enough to be used in affective detection tasks. Therefore, we developed a new methodology to create a context based affective dictionary **(c2)(c6)(p1)**. First, a new clique exploration algorithm was developed **(j1)(c7)**, which was applied afterwards on a subtitle corpora, annotated with SentiWordNet [11] valences. All these approaches are described in **Chapter 3**. In the future, this resource could be used as a dictionary for the multi-modal affect detector, described in **Chapter 2**.

On the System part, **Chapter 4** does a brief description of the protocol, formalised as a Wizard of Oz scenario, used to collect interaction data from a storytelling environment. Technical aspects of the experiment are described as well, by presenting the Online Annotation Toolkit (OAK) **(c1)(d1)**. Moreover, several psychological results are presented to support our hypothesis, that the interaction between a child and an avatar has similar characteristics with the interaction between a child and an adult in video conference mode **(c4)(c5)**.

Chapter 5 presents the architecture of the MyBlock and AgentSlang projects **(c8)**. These are systems that allow easy modelling of component based design for building agents that deal with rich feedback data. Moreover, for the knowledge extraction part, the Syn!bad library is described.

In future, the interaction model computed from the data collected with the OAK platform would be compiled into a reactive model, which will be integrated into AgentSlang. The multi-modal emotion detection algorithm and the contextualized affect dictionary would be integrated as well into the same platform.

Publication list

Journal Papers

(j1) **O. Şerban**, A. Pauchet, A. Rogozan, and J-P. Pecuchet. DDMCE: a Dynamic Distributable Maximal Clique Algorithm, *Special Issue of International Journal of Social Network Mining.* 2013. (submitted)

Conferences

(c1) **O. Şerban**, A. Pauchet, A. Bersoult, and E. Chanoni. Tell me a story: A comparative study of child interaction with virtual characters and adults, *Thirteenth International Conference on Intelligent Virtual Agents (IVA).* 2013. (submitted)

(c2) **O. Şerban**, A. Pauchet, A. Rogozan, and J-P. Pecuchet. Modelling context to solve conflicts in SentiWordNet, *The fifth biannual Humaine Association Confer-*

ence on Affective Computing and Intelligent Interaction (ACII). 2013. (submitted)

(c3) **O. Şerban**, G. Castellano, A. Pauchet, A. Rogozan, and J-P. Pecuchet. Fusion of Smile, Valence and NGram features for automatic affect detection, *The fifth biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*. 2013. (submitted)

(c4) A. Pauchet, F. Rioult, E. Chanoni, Z. Ales and **O. Şerban**. Advances on Dialogue Modelling Interactive Narration Requires Prominent Interaction and Emotion, In Joaquim Filipe and Ana Fred, editors, *Proceedings of the 5th International Conference on Agents and Artificial Intelligence*, volume 1, pages 527–530, SciTePress, 2013.

(c5) A. Pauchet, F. Rioult, E. Chanoni, Z. Ales and **O. Şerban**. Modélisation de dialogues narratifs pour la conception d'un ACA narrateur, *Proceedings of the WACAI 2012 - Workshop Affect, Compagnon Artificiel, Interaction*, 8 pages, 2012.

(c6) **O. Şerban**, A. Pauchet, A. Rogozan and J-P. Pecuchet. Semantic Propagation on Contextonyms using SentiWordNet, *Proceedings of the WACAI 2012 - Workshop Affect, Compagnon Artificiel, Interaction*, 7 pages, 2012.

(c7) **O. Şerban**, A. Pauchet, A. Rogozan and J-P. Pecuchet. DDMCE : recherche de cliques maximales dans des graphes dynamiques de grande taille, *Proceedings of the 3ième Journée thématique : Fouille de grands graphes*, 5 pages, 2012.

(c8) Z. Ales, G. Dubuisson Duplessis, **O. Şerban** and A. Pauchet. A Methodology to Design Human-Like Embodied Conversational Agents, *Proceedings of the 1st International Workshop on Human-Agent Interaction Design and Models*, pages 34-49, 2012.

(c9) **O. Şerban**, A. Pauchet, and H.F. Pop. Recognizing emotions in short text. In Joaquim Filipe and Ana Fred, editors, *Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, volume 1, pages 477–480. SciTePress, 2012.

Demo

(d1) **O. Şerban** and A. Pauchet. OAK: The Online Annotation Kit, *Proceedings of the WACAI 2012 - Workshop Affect, Compagnon Artificiel, Interaction*, 2 pages, 2012.

Poster

(p1) **O. Şerban**, Adding affective states to contextonyms, *International Workshop on Voice and Speech Processing in Social Interactions*, Glasgow, UK, 2011

Summary

Human-Computer Interaction migrates from the classic perspective to a more natural environment, where humans are able to use natural language to exchange knowledge with a computer. In order to fully "understand" the human's intentions, the computer should be able to detect emotions and reply accordingly. This thesis focuses on several issues regarding the human affects, from various detection techniques to their integration into a Distributed Interactive System.

Emotions are a fuzzy concept and their perception across human individuals may vary as well. Therefore, this makes the detection problem very difficult for a computer. From the affect detection perspective, we proposed three different approaches: an emotion detection method based on Self Organizing Maps, a valence classifier based on multi-modal features and Support Vector Machines and a technique to resolve conflicts into a well known affective dictionary (SentiWordNet). Moreover, from the system integration perspective, two issues are approached: a Wizard of Oz experiment in a children storytelling environment and an architecture for a Distributed Interactive System.

The first detection method is based on neural network model, the Self Organizing Maps, which is easy to train, but very versatile for fuzzy classification. This method works only with textual data and it uses also an Latent Semantic Analyser (LSA) feature extraction algorithm with large dictionaries as support vectors. The issue is approached as a Statistical Machine Learning problem and the validation is conducted on a well known corpus for semantic affect recognition: SemEval 2007, task 14. This experiment leads to a classification model that provides a good balance between precision and recall, for the given corpus.

We continue on the same Machine Learning perspective, by conducting a multi-modal classification study on a Youtube corpus. The smile, as a gesture feature, is fused with several other features extracted from textual data. We study the influence of smile across different configurations, with a two level linear Support Vector Machine. This offers the possibility to study in more details the classification process and therefore, we obtain the best results for the proposed corpus.

In the field of Emotion Detection the focus is mainly on two aspects: finding the

best detection algorithms and building better affective dictionaries. Whereas the first problem is tackled by the algorithms previously presented, we also focus on the second issue as well. We are decreasing the number of inconsistencies of an existing linguistic resource, the SentiWordNet dictionary, by introducing context. This is modelled as a context graph (*contextonyms*), built using a subtitle database. By applying our technique, we managed to obtain a low conflict rate, while the size of the dictionary is preserved. Our final goal is to obtain a large affective dictionary that can be used for emotion classification tasks. Decreasing the number of inconsistencies in this dictionary would directly improve the precision of the method using it.

The contextonyms are cliques in a graph of word co-occurrences. Therefore, these represent a strong semantic relation between the terms, similar to synonymic relation. The clique extraction algorithm used for this purpose was designed for building the contextonym graph, since none of the existing algorithms could handle large and dynamic graph structures. Our algorithm, the Dynamic Distributable Maximal Clique Exploration Algorithm (DDMCE), was successfully validated on various random generated databases.

From the system integration perspective, the problem of Child-Machine interaction is tackled through a storytelling environment. From the psychological perspective, this experiment is a validation of the interactive engagement between a child and a virtual character. The engineering aspects of this experiment lead to the development of a new Wizard of Oz platform (OAK), that allows online annotation of the data. Moreover, this environment helps on designing and building new reactive dialogue models, which can be integrated into our future system.

The second aspect of system integration is tackled by building a new architecture for a Distributed Interactive System. This is constructed around the idea of component based design, where the structure of the component is simple enough to allow the integration of any existing algorithm. The proposed platform currently offers several components for knowledge extraction, reactive dialogue management and affective feedback detection, among other classic components (i.e. Automatic Speech Recognition, Text to Speech). Moreover, all the algorithms previously presented can be integrated into this platform as different components.

Conclusion

The issues covered by this thesis are linked to various Affective Computing and Interaction problems. The computer should be able to detect emotions and reply accordingly.

The Affective Computing part is represented by several Detection Algorithms: a text-based emotion classification method using Self Organizing Maps and a valence classifier based on multi-modal features and Support Vector Machines. Moreover, due to an observation made while developing the classification algorithms and since one representative dictionary for sentiment analysis (SentiWordNet) carries a large number of conflicts, we also proposed a method to solve these conflicts. From the Interaction perspective, this thesis approaches two issues: an experiment created for collecting rich interactive data, in a story telling environment and an architecture for a Distributed Interactive System.

Detection of User's Affective Feedback

The field of Emotion Detection focuses on two major aspects: creating better detection algorithms and building more accurate affective dictionaries. We approached both of these issues in this thesis.

Our first experiment is based on a Self Organizing Map classifier, which is easy to train, but very versatile for fuzzy classification. We used a feature extraction using a Latent Semantic Analysis on text, which served as support for our classifier. The approach has been validated on a well known corpus for semantic affect recognition: SemEval 2007, task 14. For this purpose, we managed to obtain a model that provides a good balance between precision and recall, for the given corpus.

This first approach uses only text to extract features. Several recent studies in Affective Computing propose multi-modal approaches. Our second experiment is conducted as a multi-modal classification study on a Youtube corpus. The smile, as a gesture feature, is fused with several other features extracted from textual data. For this purpose, we generate different feature configurations to study the smile influence. These features

are used with a two level linear Support Vector Machine, which offers the possibility to study in more details the classification process. On the Youtube corpus, by using this approach, we managed to obtain the best results, compared to the original Morency et al. [126] approach, with a method that is fast enough for an interactive system.

Several issues regarding the classification precision and recall for Affective Computing are linked with the dictionaries used. These are either manually constructed with a size too small to cover all the semantic cases or very large in size but carrying a large number of internal conflicts. Decreasing the number of inconsistencies in a dictionary directly improves the precision of the method using it. We proposed to decrease the number of inconsistencies of an existing dictionary (SentiWordNet) by introducing context. The context is modelled as a contextonym graph, built using a subtitle database. We managed to obtain a low conflict rate, while the size of the dictionary is preserved. By using our method, our goal is to obtain a large contextualised affective dictionary that can be used for emotion classification tasks.

The contextonym is modelled as a strong semantic relation between the terms, similar to synonyms. In fact, these are cliques in a graph of word co-occurrences. Since none of the existing algorithms could handle large and dynamic graph structures, the clique extraction algorithm used for this purpose was designed for building the contextonym graph. Our algorithm, the Dynamic Distributable Maximal Clique Algorithm (DDMCE), was successfully validated on various random generated databases. One of the strong points of this algorithms is the it addresses the issue of processing in a distributable way large and dynamic data simultaneously. Moreover, the most important validation is represented by the ability to successfully process the data needed to generate our contextonyms model.

Affective Interactive Systems

From the interaction perspective, in order to obtain a rich corpus for the problem of Child-Machine interaction, we created an innovative storytelling environment. From the psychological perspective, this experiment is a validation of the interactive engagement between a child and a virtual character. We measured the difference between a setup having the virtual character as a narrator or a psychologist in video conference mode, by using various communicative features. The only difference, we observed, between the two scenarios is in the communication modality. This experiment also lead to the development of a new Wizard of Oz platform (OAK), that allows online annotation of the data. This environment allows the design of various reactive dialogue models, which can be tested and integrated into our future system.

The final aspect of this thesis is the proposition of a new architecture for a Distributed Interactive System. By using a component based design approach, we model a component structure that is light and simple enough to allow the integration of any existing algorithm. We propose several components for knowledge extraction (Syn!bad), reactive dialogue management and affective feedback detection, among other classic components (i.e. Automatic Speech Recognition, Text to Speech). This platform in-

tends to be the foundation for several affect detection algorithms, starting with all the algorithms previously presented.

Future work

From the affect detection perspective, the multi-modal approaches should be investigated further. Our experience with the ACAMODIA Project showed that in practice, no modality carries more importance than the others. In the storytelling experiment, the feedback is sometimes recovered from speech, gestures, postures, smiles or eye-gazing. In almost all the situations, the feedback is recovered only from one modality, while the other are missing or occluded. For example an occlusion phenomenon in speech appears when a noise covers the dialogue making the word recognition impossible, while an occlusion phenomenon in gestures appears for instance in a situation where the face is covered by the user's hand. Working with children makes this problem more difficult because all these issues appear with a high frequency.

From the dictionary conflict resolution, our current approach uses subtitles to compile a non-formal, dialogue style, linguistic model. In future, for different language styles a corresponding context graph could be constructed. This could improve the detection results for more formal environments. In the end, a large context structure can be compiled, with words clustered by style, domain and part-of-speech. Each of these propositions raises issues related to the Big Data Processing domain, which currently have not been discussed by the thesis.

The AgentSlang system offers a good foundation for building Affective Interactive Systems. Nevertheless, the number of usable components needs to increase. Currently, we propose at least one element to solve each representative problem in our design flow, but more need to be integrated. The dialogue management module, one of the critical parts of our proposition, is currently a reactive approach. The state based model, developed for the ACAMODIA Project needs to be integrated as well. Moreover, a more complex management based on Dialogue Games Theory [51], like the one we proposed for Ales et al. [6], could be developed.

Our final proposition, MyBlock has been validated only in terms of performance against a similar system, the SEMAINE platform. The technical validation is only the first step and a more extensive study of integration and acceptability needs to be conducted with human users. For this purpose, we propose two directions: *a*) a demonstrator having only Affective Feedback detection and synthesis capabilities, with a basic reactive dialogue management, similar to the Sensitive Artificial Listener proposed by SEMAINE *b*) a system having complex dialogue management based on ACAMODIA Project or Dialogue Games Theory [6, 51]. The first direction allows to compare our platform with SEMAINE, using the same scenario. Whereas, the second approach is more complex and integrates all the components currently developed (i.e. Affective Feedback Detection, Knowledge Extraction, Dialogue Management). This study can be conducted in a storytelling environment, similar to ACAMODIA, or a scenario dependent task, such as *"How was your day ?"* proposed by the Companions Project.

Bibliography

[1] Pauchet A., Rioult F., Chanoni E., Ales Z., and Şerban O. Advances on dialogue modelling interactive narration requires prominent interaction and emotion. In Joaquim Filipe and Ana Fred, editors, *Proceedings of the 5th International Conference on Agents and Artificial Intelligence*, volume 1, pages 527–530. SciTePress, 2013.

[2] Steven Abney. Parsing by chunks. *Principle-based parsing*, 44:257–278, 1991.

[3] Amel Achour, Marc Le Tallec, Sébastien Saint-Aimé, Brigitte Le Pévédic, Jeanne Villaneau, J-Y Antoine, and Dominique Duhaut. Emotirob: from understanding to cognitive interaction. In *Mechatronics and Automation, 2008. ICMA 2008. IEEE International Conference on*, pages 369–374. IEEE, 2008.

[4] C.C. Aggarwal, J.B. Orlin, and R.P. Tai. Optimized crossover for the independent set problem. *Operations Research*, 45(2):226–234, 1997.

[5] EA Akkoyunlu. The enumeration of maximal cliques of large graphs. *SIAM Journal on Computing*, 2:1, 1973.

[6] Z. Ales, G. Dubuisson Duplessis, O. Şerban, and A. Pauchet. A methodology to design human-like embodied conversational agents. *Proceedings of the 1st International Workshop on Human-Agent Interaction Design and Models*, pages 34–49, 2012.

[7] V. Alfred. Algorithms for finding patterns in strings. *Handbook of Theoretical Computer Science: Algorithms and complexity*, 1:255, 1990.

[8] Jens Allwood. An activity based approach to pragmatics. *Abduction, belief and context in dialogue: Studies in computational pragmatics*, pages 47–80, 2000.

[9] C.O. Alm, D. Roth, and R. Sproat. Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 579–586. Association for Computational Linguistics, 2005.

[10] Apple. Siri. `http://www.apple.com/ios/siri/`, March 2013.

[11] S. Baccianella, A. Esuli, and F. Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Seventh conference on International Language Resources and Evaluation, Malta. Retrieved May*, volume 25, page 2010, 2010.

[12] E. Balas and W. Niehaus. Optimized crossover-based genetic algorithms for the maximum cardinality and maximum weight clique problems. *Journal of Heuristics*, 4(2):107–122, 1998.

[13] O. Bälter, O. Engwall, A.-M. Öster, and H. Kjellström. Wizard-of-oz test of artur: a computer-based speech training system with articulation correction. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pages 36–43. ACM, 2005.

[14] A. Baylor and J. Ryu. The api (agent persona instrument) for assessing pedagogical agent persona. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, volume 2003, pages 448–451, 2003.

[15] R. Beale and C. Creed. Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies*, 67(9):755–776, 2009.

[16] C. Benzmüller, H. Horacek, I. Kruijff-Korbayová, H. Lesourd, M. Schiller, and M. Wolska. Diawoz-ii–a tool for wizard-of-oz experiments in mathematics. In *KI 2006: Advances in Artificial Intelligence*, pages 159–173. Springer, 2007.

[17] A. Bersoult. Les enfants de 7 ans interagissent-ils avec un agent virtual comme evec un adulte ? un situation de narration. Master's thesis, Universite de Rouen, 2012.

[18] Timothy W Bickmore and Rosalind W Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.

[19] S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. A. Reis, and J. Reynar. Building a sentiment summarizer for local service reviews. In *WWW Workshop on NLP in the Information Explosion Era*, 2008.

[20] R.C. Bolles. Robust feature matching through maximal cliques. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 182, page 140, 1979.

[21] I.M. Bomze, M. Budinich, P.M. Pardalos, and M. Pelillo. The maximum clique problem. *Handbook of combinatorial optimization*, 4(1):1–74, 1999.

[22] I. Borg and P.J.F. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Verlag, 2005.

[23] U. Brandes and T. Erlebach. *Network analysis: Methodological foundations*, volume 3418. Springer Verlag, 2005.

[24] C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM*, 16(9):575–577, 1973.

[25] H. Bunke and K. Shearer. A graph distance metric based on the maximal common subgraph. *Pattern recognition letters*, 19(3):255–259, 1998.

[26] D. C. Burnett, M. R. Walker, and A. Hunt. Speech synthesis markup language (ssml) version 1.0. `http://www.w3.org/TR/speech-synthesis/`, September 2009.

[27] C. C. Russell and M. K. Odell. Soundex. `http://www.archives.gov/research/census/soundex.html`, March 2013.

[28] R.A. Calvo and S. D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, pages 18–37, 2010.

[29] R. Carraghan and P.M. Pardalos. An exact algorithm for the maximum clique problem. *Operations Research Letters*, 9(6):375–382, 1990.

[30] J. Cassell. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. *Embodied conversational agents*, pages 1–27, 2000.

[31] Marc Cavazza, Raul Santos de la Camara, and Markku Turunen. How was your day?: a companion eca. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1629–1630. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[32] F. Cazals and C. Karande. A note on the problem of reporting maximal cliques. *Theoretical Computer Science*, 407(1-3):564–568, 2008.

[33] CERN. The colt distribution: Open source libraries for high performance scientific and technical computing in java, April 2012. CERN - European Organization for Nuclear Research.

[34] C. Chang, S. Chatterjee, and P.R. Kube. On an analysis of static occlusion in stereo vision. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 722–723. IEEE, 1991.

[35] F. R. Chaumartin. Upar7: A knowledge-based system for headline sentiment tagging. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 422–425, 2007.

[36] W.J. Christmas, J. Kittler, and M. Petrou. Structural matching in computer vision using probabilistic relaxation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(8):749–764, 1995.

[37] Herbert H Clark. *Using language*, volume 4. Cambridge University Press Cambridge, 1996.

[38] CMLabs. Psyclone. `http://www.mindmakers.org/projects/psyclone/`, 2007.

[39] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *The Journal of Machine Learning Research*, 12:2493–2537, 2011.

[40] Matthieu Courgeon, Jean-Claude Martin, and Christian Jacquemin. Marc: a multimodal affective and reactive character. In *Proceedings of the 1st Workshop on AFFective Interaction in Natural Environments*, 2008.

[41] Roddy Cowie. Describing the forms of emotional colouring that pervade everyday life. *Oxford handbook of philosophy of emotion. Oxford University Press, London*, pages 63–94, 2010.

[42] Roddy Cowie, Gary McKeown, and Ellen Douglas-Cowie. Tracing emotion: an overview. *International Journal of Synthetic Emotions (IJSE)*, 3(1):1–17, 2012.

[43] R. Cox. Regular expression matching can be simple and fast (but is slow in java, perl, php, python, ruby, ...). `http://swtch.com/~rsc/regexp/regexp1.html`, January 2007.

[44] T. Danisman and A. Alpkocak. Feeler: Emotion classification of text using vector space model. In *AISB 2008 Convention Communication, Interaction and Social Intelligence*, volume 1, page 53, 2008.

[45] C. Darwin, P. Ekman, and P. Prodger. *The expression of the emotions in man and animals*. Oxford University Press, USA, 2002.

[46] Nikolaos Dimakis, John K Soldatos, Lazaros Polymenakos, Pascal Fleury, Jan Curín, and Jan Kleindienst. Integrated development of context-aware applications in smart spaces. *Pervasive Computing, IEEE*, 7(4):71–79, 2008.

[47] S.K. D'Mello, S.D. Craig, J. Sullins, and A.C. Graesser. Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education*, 16(1):3–28, 2006.

[48] E. Douglas-Cowie, R. Cowie, C. Cox, N. Amier, and D. Heylen. The sensitive artificial listner: an induction technique for generating emotionally coloured conversation. *Preceedings of Workshop on Corpora for Research on Emotion and Affect*, 2008.

[49] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. Mcrorie, J.C. Martin, L. Devillers, S. Abrilian, A. Batliner, et al. The HUMAINE database: addressing the collection and annotation of naturalistic and induced emotional data. *Affective computing and intelligent interaction*, pages 488–500, 2007.

[50] N. Du, B. Wu, L. Xu, B. Wang, and X. Pei. A parallel algorithm for enumerating all maximal cliques in complex network. In *Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on*, pages 320–324. IEEE, 2006.

[51] G. Dubuisson Duplessis, N. Chaignaud, J-Ph. Kotowicz, A. Pauchet, and J-P. Pécuchet. Empirical specification of dialogue games for an interactive agent. In Y. Demazeau, T. Ishida, J.M. Corchado, and J. Bajo, editors, *Proceedings of the 11th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS)*, volume LNAI 7879, pages 49–60, May 2013.

[52] Ken Ducatel, Marc Bogdanowicz, Fabiana Scapolo, Jos Leijten, and Jean-Claude Burgelman. *Scenarios for ambient intelligence in 2010*. Office for official publications of the European Communities, 2001.

[53] S.T. Dumais. Latent semantic analysis. *Annual Review of Information Science and Technology*, 38(1):188–230, 2004.

[54] P. Ekman et al. *Universals and cultural differences in facial expressions of emotion*. University of Nebraska Press, 1971.

[55] P. Ekman et al. An argument for basic emotions. *Cognition and emotion*, 6(3/4):169–200, 1992.

[56] P. Ekman and W.V. Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1):49–98, 1969.

[57] P. Ekman and W.V. Friesen. Facial Action Coding System (FACS): A technique for the measurement of facial action. *Palo Alto, CA: Consulting*, 1978.

[58] P. Ekman and E.L. Rosenberg. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA, 2005.

[59] Paul Ekman. Basic emotions. *Handbook of cognition and emotion*, 98:45–60, 1999.

[60] F. Eyben, M. Wöllmer, and B. Schuller. Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the international conference on Multimedia*, pages 1459–1462. ACM, 2010.

[61] C. Faloutsos and D.W. Oard. A survey of information retrieval and filtering methods. *University of Maryland Computer Science Department*, 1:23, 1998.

[62] C. Fellbaum et al. Wordnet and wordnets. *Encyclopedia of Language and Linguistics*, pages 665–670, 2005.

[63] Al Geist, Adam Beguelin, Jack Dongarra, Weicheng Jiang, Robert Manchek, and Vaidy Sunderam. *PVM: Parallel virtual machine: a users\'guide and tutorial for networked parallel computing*. MIT Press, 1994.

[64] Rodolphe Gelin, Christophe d'Alessandro, Q. Anh Le, Olivier Deroo, David Doukhan, Jean-Claude Martin, Catherine Pelachaud, Albert Rilliard, and Sophie Rosset. Towards a storytelling humanoid robot. In *AAAI Fall Symposium Series*, 2010.

[65] A.M. Glenberg, D. Havas, R. Becker, and M. Rinck. Grounding language in bodily states. *Grounding cognition: the role of perception and action in memory, language, and thinking*, page 115, 2005.

[66] N. Godbole, M. Srinivasaiah, and S. Skiena. Large-scale sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, volume 2, 2007.

[67] Google. Protocol buffers. `https://developers.google.com/protocol-buffers/`, 2012.

[68] Google. Speech recognition api. `http://developer.android.com/reference/android/speech/package-summary.html`, March 2013.

[69] D. Greene, M. Parnas, and F. Yao. Multi-index hashing for information retrieval. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 722–731. IEEE, 1994.

[70] A. Grosso, M. Locatelli, and F.D. Croce. Combining swaps and node weights in an adaptive greedy approach for the maximum clique problem. *Journal of Heuristics*, 10(2):135–152, 2004.

[71] A. Grosso, M. Locatelli, and W. Pullan. Simple ingredients leading to very efficient heuristics for the maximum clique problem. *Journal of Heuristics*, 14(6):587–612, 2008.

[72] O. Grynszpan, J.-C. Martin, and J. Nadel. Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal of Human-Computer Studies*, 66(8):628–639, 2008.

[73] H. Gunes and B. Schuller. Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, 2012.

[74] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[75] J. Han, M. Jo, S. Park, and S. Kim. The educational use of home robots for children. In *Robot and Human Interactive Communication*, pages 378–383. IEEE, 2005.

[76] E.R. Hancock and J. Kittler. Edge-labeling using dictionary-based relaxation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(2):165–181, 1990.

[77] Gerd Herzog, Alassane Ndiaye, Stefan Merten, Heinz Kirchmann, Tilman Becker, and Peter Poller. Large-scale software integration for spoken language and multimodal dialog systems. *Natural Language Engineering*, 10(3-4):283–305, 2004.

[78] D. Heylen, S. Kopp, S. Marsella, C. Pelachaud, and H. Vilhjalmsson. Why conversational agents do what they do? functional representations for generating conversational agent behavior. In *Proceedings of the First Functional Markup Language Workshop*, 2008.

[79] P. Hintjens. *Zeromq: Messaging for Many Applications*. O'Reilly Media, 2013.

[80] R. Horaud and T. Skordas. Stereo correspondence through feature grouping and maximal cliques. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(11):1168–1180, 1989.

[81] Hung-Hsuan Huang, Aleksandra Cerekovic, Kateryna Tarasenko, Vjekoslav Levacic, Goranka Zoric, Igor S Pandzic, Yukiko Nakano, and Toyoaki Nishida. Integrating embodied conversational agent components with a generic framework. *Multiagent and Grid Systems*, 4(4):371–386, 2008.

[82] Hung-Hsuan Huang, Aleksandra Cerekovic, Kateryna Tarasenko, Vjekoslav Levacic, Goranka Zoric, Margus Treumuth, Igor S Pandzic, Yukiko Nakano, and Toyoaki Nishida. An agent based multicultural user interface in a customer service application. In *The eNTERFACEâĂŹ06 international workshop on multimodal interfaces. Dubrovnik, Croatia*, 2006.

[83] Hung-Hsuan Huang, Kateryna Tarasenko, Toyoaki Nishida, Aleksandra Cerekovic, Vjekoslav Levacic, Goranka Zoric, Igor S Pandzic, and Yukiko Nakano. An agent based multicultural tour guide system with nonverbal user interface. *Journal on Multimodal User Interfaces*, 1(1):41–48, 2007.

[84] David Huggins-Daines, Mohit Kumar, Arthur Chan, Alan W Black, Mosur Ravishankar, and Alex I Rudnicky. Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 1, pages I–I. IEEE, 2006.

[85] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, pages 241–250. ACM, 2008.

[86] HUMAINE Community Portal. http://emotion-research.net/, March 2011.

[87] A. Hunt and S. McGlashan. Speech recognition grammar specification version 1.0. http://www.w3.org/TR/speech-grammar/, March 2004.

[88] Ji Hyungsuk, S. Ploux, and E. Wehrli. Lexical knowledge representation with contextonyms. In *Proceedings of the 2003 MT Summit IX*. Association for Machine Translation in the Americas, 2003.

[89] Pavel Ircing, Jan Romportl, and Zdenek Loose. Audiovisual interface for czech spoken dialogue system. In *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, pages 526–529. IEEE, 2010.

[90] iSpeech Inc. ispeech tts service. http://www.ispeech.org/, March 2013.

[91] K. Jaber and R.A. Nur'Aini Abdul Rashid. The parallel maximal cliques algorithm for protein sequence clustering. *American Journal of Applied Sciences*, 6(7):1368–1372, 2009.

[92] W. James. What is an Emotion? *Mind*, 9(34):188–205, 1884.

[93] E. Jennings and L. Motyčková. A distributed algorithm for finding all maximal cliques in a network graph. *LATIN'92: 1st Latin American Symposium on Theoretical Informatics, São Paulo, Brazil, April 6-10, 1992: proceedings*, 583:281–293, 1992.

[94] H. Ji, S. Ploux, and E. Wehrli. Lexical knowledge representation with contextonyms. In *Proceedings of MT Summit IX, New Orleans, USA*. Association for Machine Translation in the Americas, 2003.

[95] Y. Ji, X. Xu, and G.D. Stormo. A graph theoretical approach for predicting common rna secondary structure motifs including pseudoknots in unaligned sequences. *Bioinformatics*, 20(10):1591–1602, 2004.

[96] D.S. Johnson and M.A. Trick. *Cliques, coloring, and satisfiability: second DIMACS implementation challenge, October 11-13, 1993*, volume 26. Amer Mathematical Society, 1996.

[97] M. Johnston, P. Baggia, D. C. Burnett, J. Carter, D. A. Dahl, G. McCobb, and D. Raggett. Emma: Extensible multimodal annotation markup language. http://www.w3.org/TR/emma/, February 2009.

[98] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19(1):61–84, 2004.

[99] R.M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, 40(4):85–103, 1972.

[100] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy. Improvements to platt's smo algorithm for svm classifier design. *Neural Computation*, 13(3):637–649, 2001.

[101] J. Klein, Y. Moon, and R. W. Picard. This computer responds to user frustration:: Theory, design, and results. *Interacting with computers*, 14(2):119–140, 2002.

[102] S. R. Klemmer, A. K. Sinha, J. Chen, J. A Landay, N. Aboobaker, and A. Wang. Suede: a wizard of oz prototyping tool for speech user interfaces. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 1–10. ACM, 2000.

[103] I. Koch. Enumerating all connected maximal common subgraphs in two graphs. *Theoretical Computer Science*, 250(1-2):1–30, 2001.

[104] I. Koch, T. Lengauer, and E. Wanke. An algorithm for finding maximal common subtopologies in a set of protein structures. *Journal of computational biology*, 3(2):289–306, 1996.

[105] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

[106] S. Kopp, B. Krenn, S. Marsella, A. Marshall, C. Pelachaud, H. Pirker, K. Thórisson, and H. Vilhjálmsson. Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent Virtual Agents*, pages 205–217. Springer, 2006.

[107] Z. Kozareva, B. Navarro, S. Vázquez, and A. Montoyo. Ua-zbsa: A headline emotion classification through web information. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 334–337. Association for Computational Linguistics, 2007.

[108] H. Kozima, C. Nakagawa, and Y. Yasuda. Interactive robots for communication-care: a case-study in autism therapy. In *Robot and Human Interactive Communication*, pages 341–346, 2005.

[109] Paul Lamere, Philip Kwok, William Walker, Evandro Gouvea, Rita Singh, Bhiksha Raj, and Peter Wolf. Design of the cmu sphinx-4 decoder. In *Proceedings of the 8th European conference on speech communication and technology*, pages 1181–1184, 2003.

[110] Jean-Yves Lionel Lawson, Ahmad-Amr Al-Akkad, Jean Vanderdonckt, and Benoit Macq. An open source workbench for prototyping multimodal interactions based on off-the-shelf heterogeneous components. In *Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems*, pages 245–254. ACM, 2009.

[111] Akinobu Lee, Tatsuya Kawahara, and Kiyohiro Shikano. Julius: an open source real-time large vocabulary recognition engine. `http://julius.sourceforge.jp/`, August 2012.

[112] B. Liu. Sentiment analysis and subjectivity. *Handbook of Natural Language Processing,*, pages 627–666, 2010.

[113] H. Liu and P. Singh. ConceptNet-a practical commonsense reasoning tool-kit. *BT technology journal*, 22(4):211–226, 2004.

[114] Hugo Liu, Henry Lieberman, and Ted Selker. A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 125–132. ACM, 2003.

[115] R Duncan Luce and Albert D Perry. A method of matrix analysis of group structure. *Psychometrika*, 14(2):95–116, 1949.

[116] C. Ma, H. Prendinger, and M. Ishizuka. A chat system based on emotion estimation from text and embodied conversational messengers. *Entertainment Computing-ICEC 2005*, pages 535–538, 2005.

[117] K. Makino and T. Uno. New algorithms for enumerating all maximal cliques. *Algorithm Theory*, 9:260–272, 2004.

[118] G. Malewicz, M.H. Austern, A.J.C. Bik, J.C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski. Pregel: a system for large-scale graph processing. In *Proceedings of the 2010 international conference on Management of data*, pages 135–146. ACM, 2010.

[119] G.S. Manku, A. Jain, and A. Das Sarma. Detecting near-duplicates for web crawling. In *Proceedings of the 16th international conference on World Wide Web*, pages 141–150. ACM, 2007.

[120] T. Meltzer, C. Yanover, and Y. Weiss. Globally optimal solutions for energy minimization in stereo vision using reweighted belief propagation. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 428–435. IEEE, 2005.

[121] R. Mihalcea and D. Radev. *Graph-based natural language processing and information retrieval*. Cambridge University Press, 2011.

[122] J. Miksatko, K. H. Kipp, and M. Kipp. The persona zero-effect: Evaluating virtual character benefits on a learning task with repeated interactions. In *Intelligent Virtual Agents*, pages 475–481. Springer, 2010.

[123] G.A. Miller. WordNet: a lexical database for English. *Communications of the ACM*, 38(11):39–41, 1995.

[124] A. Mohaisen, A. Yun, and Y. Kim. Measuring the mixing time of social graphs. In *Proceedings of the 10th annual conference on Internet measurement*, pages 383–389. ACM, 2010.

[125] R. Mohr and T.C. Henderson. Arc and path consistency revisited. *Artificial intelligence*, 28(2):225–233, 1986.

[126] L. P. Morency, R. Mihalcea, and P. Doshi. Towards multimodal sentiment analysis: harvesting opinions from the web. In *Proceedings of the 13th international conference on multimodal interfaces*, pages 169–176. ACM, 2011.

[127] M. Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[128] M. Moundridou and M. Virvou. Evaluating the persona effect of an interface agent in a tutoring system. *Journal of computer assisted learning*, 18(3):253–261, 2002.

[129] C. Munteanu and M. Boldea. Mdwoz: A wizard of oz environment for dialog systems development. In *LRECâĂŹ00*. Citeseer, 2000.

[130] N.M. Nasrabadi and C.Y. Choo. Hopfield network for stereo vision correspondence. *Neural Networks, IEEE Transactions on*, 3(1):5–13, 1992.

[131] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. Textual affect sensing for sociable and expressive online communication. In *Affective Computing and Intelligent Interaction*, pages 218–229. Springer, 2007.

[132] NIST II Smart space library. `http://www.nist.gov/smartspace/`, March 2013.

[133] Nuance. Dragon speech recognition software. `http://www.nuance.com/dragon/index.htm`, March 2013.

[134] M. Ochs, J. Ollivier, B. Coic, T. Brien, and F. Majeric. Affimo: Toward an open-source system to detect affinities and emotions in userâĂŹs sentences. In *Workshop Affect, Compagnon Artificiel, Interaction (WACAI), Grenoble, France*, volume 1, 2012.

[135] M. Ochs, J. Ollivier, B. Coic, T. Brien, and F. Majeric. Affimo: Toward an open-source system to detect affinities and emotions in userâĂŹs sentences. In *Workshop Affect, Compagnon Artificiel, Interaction (WACAI), Grenoble, France*, volume 1, 2012.

[136] M. Ochs and H. Prendinger. A virtual characterâĂŹs emotional persuasiveness. In *Kansei Engineering and Emotion Research International Conference*, 2010.

[137] Jeff Orkin and Deb Roy. Automatic learning and generation of social behavior from collective human gameplay. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 385–392. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[138] C.E. Osgood, W.H. May, and M.S. Miron. *Cross-cultural universals of affective meaning.* University of Illinois Press, 1975.

[139] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum. *The measurement of meaning.* University of Illinois Press, 1971.

[140] Alexander Osherenko and Elisabeth André. Differentiated semantic analysis in lexical affect sensing. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–6. IEEE, 2009.

[141] M. Otto, R. Friesen, and D. Rösner. Message oriented middleware for flexible wizard of oz experiments in hci. In *Human-Computer Interaction. Design and Development Approaches*, pages 121–130. Springer, 2011.

[142] S. Oviatt. Talking to thimble jellies: ChildrenâĂŹs conversational speech with animated characters. In *Proc. ICSLP (Beijing, China)*, pages 67–70, 2000.

[143] Oxford Dictionary, Online Edition. http://oxforddictionaries.com/, March 2011.

[144] P. P. Leach, M. Mealling, and R. Salz. A universally unique identifier (uuid) urn namespace. `http://www.ietf.org/rfc/rfc4122.txt`, July 2005.

[145] Sathish Pammi, Marcela Charfuelan, and Marc Schröder. Multilingual voice creation toolkit for the mary tts platform. *Proc. LREC. Valettea, Malta: ELRA*, 2010.

[146] P.M. Pardalos, J. Rappe, and M.G.C. Resende. An exact parallel algorithm for the maximum clique problem. *High performance algorithms and software in nonlinear optimization*, 24:279, 1998.

[147] P.A. Pevzner, S.H. Sze, et al. Combinatorial approaches to finding subtle signals in dna sequences. In *Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology*, volume 8, pages 269–278, 2000.

[148] L. Philips. Metaphone 3. `http://www.amorphics.com/`, March 2013.

[149] Lawrence Philips. Hanging on the metaphone. *Computer Language*, 7(12 (December)), 1990.

[150] Lawrence Philips. The double metaphone search algorithm. *CC Plus Plus Users Journal*, 18(6):38–43, 2000.

[151] Rosalind W Picard and Jonathan Klein. Computers that recognise and respond to user emotion: theoretical and practical implications. *Interacting with computers*, 14(2):141–169, 2002.

[152] R.W. Picard. *Affective computing.* The MIT press, 2000.

[153] Yoann Pigné, Antoine Dutot, Frédéric Guinand, and Damien Olivier. Graphstream: A tool for bridging the gap between complex systems and dynamic graphs. *arXiv preprint arXiv:0803.2093*, 2008.

[154] J. C. Platt et al. Using analytic qp and sparseness to speed training of support vector machines. *Advances in neural information processing systems*, pages 557–563, 1999.

[155] I. Poggi, C. Pelachaud, F. Rosis, V. Carofiglio, and B. Carolis. Greta. a believable embodied conversational agent. *Multimodal intelligent information presentation*, pages 3–25, 2005.

[156] Christopher Potts. Sentiment analysis tutorial, November 2011.

[157] H. Prendinger and M. Ishizuka. The empathic companion: A character-based interface that addresses users'affective states. *Applied Artificial Intelligence*, 19(3-4):267–285, 2005.

[158] H. Prendinger, S. Mayer, J. Mori, and M. Ishizuka. Persona effect revisited. In *Intelligent Virtual Agents*, pages 283–291. Springer, 2003.

[159] K. Prepin and C. Pelachaud. Basics of intersubjectivity dynamics: Model of synchrony emergence when dialogue partners understand each other. In *Agents and Artificial Intelligence*, pages 302–318. Springer, 2013.

[160] W. Pullan and H.H. Hoos. Dynamic local search for the maximum clique problem. *Journal of Artificial Intelligence Research*, 25(1):159–185, 2006.

[161] D. Rao and D. Ravichandran. Semi-supervised polarity lexicon induction. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 675–682. Association for Computational Linguistics, 2009.

[162] V. Rieser and O. Lemon. Learning effective multimodal dialogue strategies from wizard-of-oz data: Bootstrapping and evaluation. In *Proceedings of ACL*, pages 638–646, 2008.

[163] B. Robins, P. Dickerson, and K. Dautenhahn. Robots as embodied beings - interactionally sensitive body movements in interactions among autistic children and a robot. In *Robot and Human Interactive Communication*, pages 54–59, 2005.

[164] A. Rosenfeld, R.A. Hummel, and S.W. Zucker. Scene labeling by relaxation operations. *Systems, Man and Cybernetics, IEEE Transactions on*, 6(6):420–433, 1976.

[165] J.A. Russell. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145–172, 2003.

[166] K. Ryokai, C. Vaucelle, and J. Cassell. Virtual peers as partners in storytelling and literacy learning. *Journal of computer assisted learning*, 19(2):195–208, 2003.

[167] Sébastien Saint-Aimé, Brigitte Le-Pevedic, Dominique Duhaut, and Takanori Shibata. Emotirob: companion robot project. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 919–924. IEEE, 2007.

[168] D. Salber and J. Coutaz. Applying the wizard of oz technique to the study of multimodal systems. *Human-Computer Interaction*, pages 219–230, 1993.

[169] R. Samudrala and J. Moult. A graph-theoretic algorithm for comparative modeling of protein structure1. *Journal of molecular biology*, 279(1):287–302, 1998.

[170] B. Santorini. Part-of-speech tagging guidelines for the penn treebank project (3rd revision). Technical report, University of Pennsylvania, 1990.

[171] Klaus R Scherer. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3459–3474, 2009.

[172] Klaus R Scherer, Elizabeth Clark-Polner, and Marcello Mortillaro. In the eye of the beholder? universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, 46(6):401–435, 2011.

[173] Albrecht Schmidt. Interactive context-aware systems interacting with ambient intelligence. *Ambient intelligence*, pages 159–178, 2005.

[174] M.C. Schmidt, N.F. Samatova, K. Thomas, and B.H. Park. A scalable, parallel algorithm for maximal clique enumeration. *Journal of Parallel and Distributed Computing*, 69(4):417–428, 2009.

[175] M. Schröder. The semaine api: towards a standards-based framework for building emotion-oriented systems. *Advances in Human-Computer Interaction*, 2010:2–2, 2010.

[176] M. Schröder. *The SEMAINE API: A component integration framework for a naturally interacting and emotionally competent Embodied Conversational Agent.* PhD thesis, Saarland University, 2011.

[177] M. Schröder, P. Baggia, F. Burkhardt, J.-C. Martin, C. Pelachaud, C. Peter, B. Schuller, I. Wilson, and E. Zovato. Elements of an emotionml 1.0. `http://www.w3.org/2005/Incubator/emotion/XGR-emotionml-20081120/`, November 2008.

[178] Marc Schröder, Sathish Pammi, and Oytun Türk. Multilingual mary tts participation in the blizzard challenge 2009. In *Proc. Blizzard Challenge*, volume 9, 2009.

[179] Björn Schuller, Michel Valster, Florian Eyben, Roddy Cowie, and Maja Pantic. Avec 2012: the continuous audio/visual emotion challenge. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 449–456. ACM, 2012.

[180] P. Singh, T. Lin, E. Mueller, G. Lim, T. Perkins, and W. Li Zhu. Open mind common sense: Knowledge acquisition from the general public. *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, pages 1223–1237, 2002.

[181] M. Snir, S.W. Otto, D.W. Walker, J. Dongarra, and S. Huss-Lederman. *MPI: The complete reference*. MIT press, 1995.

[182] B. Snyder, D. Bosanac, and R. Davies. *ActiveMQ in Action*. Manning Publications, 2011.

[183] Mohammad Soleymani, Maja Pantic, and Thierry Pun. Multimodal emotion recognition in response to videos. *Affective Computing, IEEE Transactions on*, 3(2):211–223, 2012.

[184] C. Solnon and S. Fenet. A study of aco capabilities for solving the maximum clique problem. *Journal of Heuristics*, 12(3):155–180, 2006.

[185] V. Stix. Finding all maximal cliques in dynamic graphs. *Computational Optimization and applications*, 27(2):173–186, 2004.

[186] Philip J Stone, Dexter C Dunphy, and Marshall S Smith. *The general inquirer: a computer approach to content analysis*. The MIT Press, 1966.

[187] C. Strapparava and R. Mihalcea. Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560. ACM, 2008.

[188] C. Strapparava and A. Valitutti. WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*, volume 4, pages 1083–1086. Citeseer, 2004.

[189] Dag Sverre Syrdal, Michael L Walters, Nuno Otero, Kheng Lee Koay, and Kerstin Dautenhahn. He knows when you are sleeping-privacy and the personal robot companion. In *Proc. Workshop Human Implications of Human-Robot Interaction, Association for the Advancement of Artificial Intelligence (AAAIâĂŹ07)*, pages 28–33, 2007.

[190] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, 2010.

[191] W. Taymans, S. Baker, A. Wingo, R. Bultje, and S. Kost. Gstreamer application development manual, 2001.

[192] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.

[193] Kristinn R Thórisson, Hrvoje Benko, Denis Abramov, Andrew Arnold, Sameer Maskey, and Aruchunan Vaseekaran. Constructionist design methodology for interactive intelligences. *AI Magazine*, 25(4):77, 2004.

[194] Erik F Tjong Kim Sang and Sabine Buchholz. Introduction to the conll-2000 shared task: Chunking. In *Proceedings of the 2nd workshop on Learning language in logic and the 4th conference on Computational natural language learning-Volume 7*, pages 127–132. Association for Computational Linguistics, 2000.

[195] Erik F Tjong Kim Sang and Fien De Meulder. Introduction to the conll-2003 shared task: Language-independent named entity recognition. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pages 142–147. Association for Computational Linguistics, 2003.

[196] E. Tomita, A. Tanaka, and H. Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theoretical Computer Science*, 363(1):28–42, 2006.

[197] S. Tsukiyama, M. Ide, H. Ariyoshi, and I. Shirakawa. A new algorithm for generating all the maximal independent sets. *SIAM Journal on Computing*, 6:505, 1977.

[198] A. Valitutti, C. Strapparava, and O. Stock. Lexical resources and semantic similarity for affective evaluative expressions generation. *Affective Computing and Intelligent Interaction*, pages 474–481, 2005.

[199] VHMsg Library. `https://confluence.ict.usc.edu/display/VHTK/VHMsg`, March 2013.

[200] C. Von Hofsten and K. Rosander. *From action to cognition*, volume 164. Elsevier Science, 2007.

[201] VoxForge. The open source acoustic model. http://www.voxforge.org/.

[202] Wolfgang Wahlster. *Verbmobil: foundations of speech-to-speech translation*. Springer verlag, 2000.

[203] Wolfgang Wahlster. *SmartKom: Foundations of Multimodal Dialogue Systems (Cognitive Technologies)*. Springer-Verlag New York, Inc., 2006.

[204] H.G. Wallbott, K.R. Scherer, et al. Emotion and economic development-Data and speculations concerning the relationship between economic factors and emotional experience. *European journal of social psychology*, 18(3):267–273, 1988.

[205] S. Whittaker, M. Walker, and J. Moore. Fish or fowl: A wizard of oz evaluation of dialogue strategies in the restaurant domain. In *Language Resources and Evaluation Conference*, 2002.

[206] Michael Widenius, David Axmark, and AB MySQL. *MySQL reference manual: documentation from the source*. O'Reilly Media, Incorporated, 2002.

[207] J. Wiebe, T. Wilson, and C. Cardie. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2):165–210, 2005.

[208] Steve Young, Gunnar Evermann, Dan Kershaw, Gareth Moore, Julian Odell, Dave Ollason, Valtcho Valtchev, and Phil Woodland. *The HTK book*, volume 3. Cambridge University Engineering Department, 2002.

[209] Li Zhang. Exploration of affect sensing from speech and metaphorical text. In *Learning by Playing. Game-Based Education System Design and Development*, pages 251–262. Springer, 2009.