deLearyous: An Interactive Application for Interpersonal Communication Training

Jeroen Wauters[1] • Frederik Van Broeckhoven[1] • Maarten Van Overveldt[1] • Koen Eneman[1] • Frederik Vaassen[2] • Walter Daelemans[2]

1 Introduction

For many people, interacting with clients or contractors is an important part of their professional activities. Employers often require those responsible for sales or customer services to partake in training activities to improve their interpersonal communication skills. Unfortunately, this sort of training requires specialized coaching and is thus expensive and time-consuming.

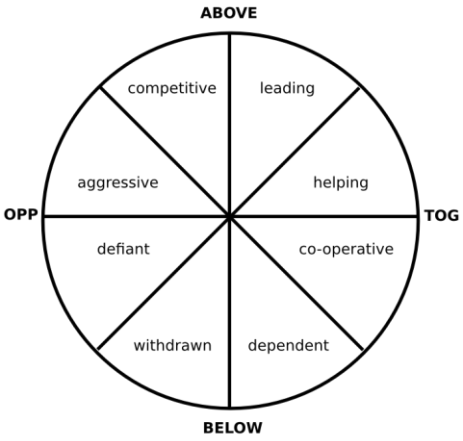

Figure 1 – The interpersonal Circumplex

Many communication training sessions are based on psychological models of communication. One of these models is the interpersonal circumplex [5,8], which models human interaction on two orthogonal axes: a vertical power axis and a horizontal love axis (figure 1). In a discussion, each conversation partner takes on a certain position on both axes, thereby influencing the position of the other. For instance, if the speaker takes on a dominant position, along the top of the power axis, he will provoke a submissive position in his conversation partner – and vice versa. By contrast, positions along the love axis are symmetric: a cooperative approach from the speaker will result in cooperation in the listener, while opposition will incite opposition [10,13].

2 Application Pipeline

With this poster presentation, we introduce some of the intermediate results of the ongoing project *deLearyous* [2], a collaboration between Group T Leuven Engineering College and the University of Antwerp. The aim of the project is to create an interactive application for interpersonal communication training, thus decreasing the need for individual coaching. Users will interact with the application through written natural language input and engage in conversation with a 3D virtual agent. They take on the role of a manager in a company, tasked with calming an employee (the virtual agent, or VA) who is disgruntled about a recent management decision to start charging money for parking privileges. The application pipeline is divided into a number of modules (figure 2), which are discussed below.

[1] e-Media Lab, Group T – International University College Leuven (Association K.U.Leuven)
Andreas Vesaliusstraat 13, 3000 Leuven, Belgium, +32(0)16 30 10 30
{jeroen.wauters, frederik.van.broeckh, maarten.van.overveldt, koen.eneman}@groept.be

[2] Computational Linguistics and Psycholinguistics Research Center, University of Antwerp
Lange Winkelstraat 40-42, 2000 Antwerpen, Belgium, +32 (0) 3 265 52 25
{frederik.vaassen, walter.daelemans}@ua.ac.be

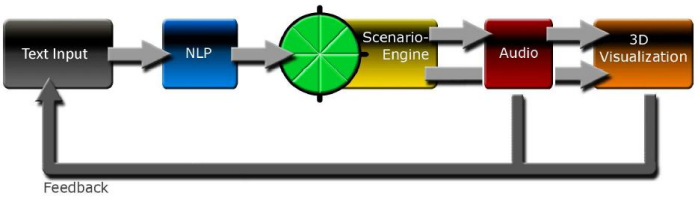

Figure 2 - Pipeline for the deLearyous application

2.1 Natural Language Processing

When the user enters a statement, his input is processed by the **natural language processing (NLP) module**. The AI of the VA is driven by the circumplex model, meaning that it uses the position of the user on the circumplex as a parameter in its decision making process. To determine the user's position, we have trained an automatic classification system [3,6] that accepts any sentence as input and provides the position of the sentence on the circumplex as output. Because the first, context-independent version of this system had a low accuracy [11], we are currently developing an upgraded classifier which uses context information to improve classification results [12]. The NLP module also identifies the keywords in the user input. It determines if keywords need to be negated and expands the keyword set through a thesaurus. This expanded set of words is then passed on, along with the predicted circumplex coordinates, to the scenario engine.

2.2 Scenario Engine

The **scenario engine** is tasked with selecting an appropriate response to the player's input on the basis of the identified keywords and the circumplex positions of both player and VA. It achieves this by feeding the input into an internal finite state machine (FSM) which models the relation between the expected statements of the user and the available responses of the VA [9]. Once the user's input sentence has been matched to a state of the FSM, a reply is selected from the available follow-up states. Additionally, the emotional state and circumplex coordinates of the VA may be updated, and this information is then sent to the audio manager.

The FSM, which forms the core of the scenario engine, has been created and refined through a series of increasingly specific user tests, ending with a wizard-of-oz test where we simulated the application by having a human actor play the VA. Care has been taken to separate application logic and scenario-specific information, so it is relatively easy for an administrator to define a new scenario and plug the corresponding FSM into the application.

2.3 Audio

The **audio module** collaborates with the 3D visualization module to output audiovisual feedback to the user. It receives input from the scenario engine and searches through its internal database of audio files, selecting the one which most closely matches the response sentence and the emotional state of the VGP.

2.4 3D Visualization

Finally, the **3D visualization module** is responsible for rendering a convincing 3D representation of the VA. It contains a database of poses and body animations, each of which is linked to a certain position on the circumplex, and possibly annotated with context information such as affirmation, negation, stress, etc. Based on the output of the scenario engine, a pose and one or more of these animations will be selected and played during playback of the audio file.

The visualization engine is also responsible for the facial animation of the VA. Facial animation is achieved by analyzing which phonemes occur in the sound file selected by the audio module and then creating an animation morph based on the head meshes for the corresponding visemes. While the results are less impressive than more advanced techniques [1,7], this approach has the benefit of simplicity and real-time generation of animations. The resulting animation is further blended with a facial representation of the VA's emotional state [4]. The head and body animations are subsequently combined into the final animation.

References

1. Cao, Y., Tien, W.C., Faloutsos, P., Pighin, F., *Expressive Speech-driven facial animation*, in ACM transactions on graphics (2005)
2. deLearyous (IWT-TETRA 090144), *Training of interpersonal communication by natural language interaction with autonomous virtual characters* (2011) http://delearyous.groept.be/
3. Ghazi, D., Inkpen, D., and Szpakowicz, S., *Hierarchical versus Flat Classification of Emotions in Text*, in Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pp. 140-146 (2010)
4. Knutson, B., Facial *Expressions of Emotion Influence Interpersonal Trait Influences*, in Journal of Nonverbal Behavior 20(3), pp.165-182 (1996)
5. Leary, T., *Interpersonal Diagnosis of Personality: Functional Theory and Methodology for Personality Evaluation*, 518p. Ronald Press, Oxford (1957)
6. Mishne, G., *Experiments with Mood Classification in Blog Posts*, in Proceedings of the 1st Workshop on Stylistic Analysis Of Text For Information Access (2005)
7. Müller, P., Kalberer, G.A., Proesmans, M., Van Gool, L., *Realistic Speech Animation Based on Observed 3D Face Dynamics*, IEE Proc. Vision, Image & Signal Processing, Vol. 152, pp.491-500 (2005)
8. Russell, J. A., *A circumplex model of affect*, in Journal of Personality and Social Psychology, 39, pp.1161-1178 (1980)
9. Traum, D.R. and Larsson, S., The Information State Approach to Dialogue Management, in Current and New Directions in Discourse and Dialogue, pp.325-353 (2007)
10. van Dijk, B., *Beïnvloed anderen, begin bij jezelf*, 118p. Thema Uitgeverij, Zaltbommel (2007)
11. Vaassen, F., and Daelemans, W., *Emotion Classification in a Serious Game for Training Communication Skills*, in Computational Linguistics in the Netherlands 2010: selected papers from the twentieth CLIN meeting (2010)
12. Vaassen, F., and Daelemans, W., *Automatic Emotion Classification for Interpersonal Communication*, in Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, pp.104-110 (2011)
13. Vermeren, P. *Rond Leiderschap: Praktijk Getoetst aan Wetenschap*, 255p. Academia Press, Gent (2009)