MISSISSIPPI (Medical Interaction Support Strategies Integrating Social Signal Processing Principles and Indicators)

M. Balsters, M. Postma Nilsenová, K. Tates, S. Brugel, E. Postma, E. Krahmer & A. Plaat
Tilburg center for Cognition and Communication
Tilburg University

Introduction
The aim of our e-Humanities project is to annotate and analyse nonverbal cues in doctor-patient interactions using available (semi-)automatic annotation techniques. The nonverbal communication skills of health professionals, which involve both the ability to decode and encode cues correctly, affect the outcome of their interaction with patients, most importantly, patient satisfaction, understanding, recall and compliance/adherence to instructions (Schellinger et al., 2003; Flocke & Stange, 2004; Hentry et al., 2011). Research has shown that, quite alarmingly, patients tend to misunderstand or forget between 20%-75% of the information they receive (Jansen, 2009), which results in additional hospital time and treatments for side effects caused, for example, by patients’ errors in taking medication. Experimental studies strongly suggest that processing of medical information can be improved by means of enhanced (tailored) communication: communication where the patient’s emotions, attitudes and cognitive resources are taken into account (van der Meulen et al., 2008). However, existing approaches to analysing medical interviews make little use of the methods and computational tools available in the field of nonverbal research and social signal processing.

In our project, we make use of existing annotation schemes to identify in a systematic way the nonverbal cues sent off both by doctors and by patients, that have an impact on relevant patient outcomes. The project is unique in its use of the e-Humanities approach to an analysis of an existing corpus of domain-specific interactions, provided by the Netherlands Institute for Health Services Research (NIVEL). Furthermore, it combines the research findings from two different fields, health communication and computational processing of nonverbal cues, which can theoretically benefit from each other.

Background
Long-standing research in the field of medical communication has shown that the way health practitioners communicate with patients has an impact on a number of important patient outcomes (Laine et al., 1996; Tates, 2001; Zandbelt et al., 2004; Makoul, 2001; Roter & Hall, 2006; Pawlikowska et al., 2012). These outcomes, such as quality of the diagnosis, patient satisfaction, reported enhanced physical and emotional health, understanding of medical information, recall and compliance/adherence to instructions, better performance in daily activities, as well as improvements in markers of disease (e.g., blood pressure), can be improved if the physician’s communicative skills improve (Stewart, 1995; Roter & Hall, 2006). In fact, next to expert knowledge, the communicative process is considered to be one of the two fundamental components of medical care (Roter & Hall, 2006). Therefore, the majority of medical programs recognise the need for a communication training (Simpson, Buckman, & Stewart, 1991) and employ various teaching methods to help medical students improve their communicative skills.
The core skills needed for a successful doctor-patient consultation are typically formulated on the basis of the Three Function Model of Medical Interviewing (Lazare, Putnam, & Lipkin, 1995; Cohen-Cole, 1991), further developed in the Macy Model of Doctor-Patient communication (Kalet et al., 2004). From a communicative perspective, the focus lies on the following skill areas:

1. Data gathering techniques that involve active listening, verbal and nonverbal encouragement, non-interruptions, open-ended questions and acknowledgments;
2. Eliciting and understanding patient’s perspective and relationship building by recognizing, acknowledging and responding to patient’s concerns, emotions and nonverbal cues, as well as by using various activation strategies;
3. Patient education by offering relevant information in language the patient can understand and by asking repeatedly for patient’s understanding.

For purposes of evaluation and training of medical professionals, their communicative performance is measured by means of implementing the Three Function Model of Medical Interviewing in the Roter Interaction Analysis System (RIAS). The RIAS has been originally developed by Debra Roter and her collaborators in the 1970s (Roter, 1999; Roter & Larson, 2002; Roter & Hall, 2006) and has since been applied in many medical domains, including oncology (Ong et al., 1998; Siminoff et al., 2000), obstetrics and pediatrics (Van Dulmen & Bensing, 2000; Wissow, Roter, & Wilson, 1994), telemedicine (Agha, Roter, & Schapira, 2009), gynecology (Roter et al., 1999), hypertension (Roter & Ewart, 1992) and diabetes (Van Dulmen, Verhaak, & Bilo, 1997). The system offers a method of coding medical interactions by categorising each utterance in the dialogue into one of the 41 mutually exclusive categories (e.g., ‘empathy statements’, ‘showing approval’, ‘showing agreement or understanding’). The categories are defined both in terms of the intended effect on the addressee (the patient) and their linguistic properties (e.g., ‘asks closed-ended questions: medical condition’ or ‘empathy statements’).

Among the advantages of the coding system are its high inter-coder reliability (Roter & Larson, 2002), availability in other languages apart from English (Brink-Muinen, Verhaak, & Bensing, 1999) and, most importantly, its predictive and concurrent validity with respect to a number of patient outcomes, such as patient and physician satisfaction, patient recall, and improved emotional health (Roter & Larson, 2002). The RIAS, however, also has a number of weaknesses, both from an applied perspective, as well as in terms of its theoretical grounding.

For the purposes of training, an obvious disadvantage of the RIAS lies in the relatively time-consuming coding of video-taped interviews. In a study examining the benefits of training medical professionals with video feedback based on the RIAS, Roter et al. (2004) estimated that the coding involves approximately three times the length of the interview and needs to be done by experienced coders (ca 100 - 110 hours of training). As suggested by Roter et al., what is actually needed is an interactive training module with on-line just in time feedback but such a module would require automatic recognition of relevant action units with a situation-dependent syntactic and semantic analysis of the utterances.

Designing an automatic annotation system for the RIAS seems to be difficult because, theoretically speaking, the system offers a one-dimensional analysis - there is no internal structure and the tags are
mutually exclusive, so that each utterance can only be classified into one category. However, dialogues are typically multifunctional: As shown in both qualitative and quantitative studies of human dialogue behaviour (Bunt, 2007; Bunt, 2009), the multi-functionality of utterances on average amounts to 4-5 functions and is even greater when nonverbal behaviour is taken into consideration. In fact, in computational studies of dialogue taxonomy, one of the basic requirements on the proposed system is that it should support the multifunctional analysis of dialogue (Allan & Core, 1997; Larsson, 1998; Popescu-Belis, 2005; Bunt, 2009).

Another disadvantage of the RIAS is the under-representation of the role of emotions in the doctor-patient interaction. Emotions have an important impact on medical interactions and their recognition and acknowledgment form an integral part of the Three Function Model of Medical Interviewing, in terms of strengthening the relationship with the patient (Shields et al., 2005), understanding the patient’s perspective (Butow et al., 2002) and patient education/enabledness (Zachariae et al., 2003; Pawlikowska et al., 2012). For patients, an illness is not a purely medical issue as in the course of the treatment they need to deal with various levels of uncertainty and emotional anxiety (Bensing, 1991; Hulsman, 1998). With respect to emotion detection, a fundamental drawback of the RIAS is its emphasis on verbalisations, which makes it less suitable to code emotional cues (Zimmermann et al., 2011). As noted by Bensing, Verheul, & Van Dulmen (2008), patients seldom verbalise their emotions explicitly in the interaction with a medical professional. In fact, psychological distress appears to be negatively correlated with an explicit (verbal) expression of concerns (Greenley, Young & Schoenherr, 1982; Heaven & Maguire, 1997), but can be detected by means of nonverbal cues (Uitterhoeve et al., 2009).

In order to address the important problem of emotional cue recognition, the Verona Coding Definitions of Emotional Sequences (VR-CoDES) approach has been developed (Zimmermann et al., 2011), consisting of a manual for Cue/Concern expression coding (CC) and a manual for health provider responses (P). In the system, ‘concerns’ are explicit verbalisations of an unpleasant emotion, while ‘cues’ are unclear verbal or nonverbal hints to the emotion, such as sighs, crying, vague ambiguous descriptions, repetitions etc. However, from the point of view of nonverbal cue, the system is quite limited in that it mostly ignores the rich information available in the face, body posture and vocal properties of the speaker. In fact, among the seven types of cues defined by the system, only one explicitly refers to nonverbal expressions and does not address them in any detail. This is rather striking, given the plethora of research that shows the importance of nonverbal cues: Information about speakers’ emotions is encoded in specific facial expression configurations (Mendolia, 2007) and vocal patterns (Juslin & Laukka, 2003; Haskard et al., 2008); gestures, among else, serve to embody information about bodily experience of pain (Rowbotham et al., 2012). In comparison, the Ethological Coding System for Interviews (ECSI), used to score visual nonverbal behaviour in clinical interactions with psychiatric patients (Troisi, 1999; Troisi et al., 2007), differentiates among 37 different behaviour patterns (mostly facial expressions and hand movements) and offers the means to interpret them as one of the seven behavioural categories (e.g., in terms of affiliation, relaxation or displacement).

The CC-scheme also does not distinguish between different dimensions of emotions, in particular, their valence and arousal, and their potency (Goubeek & Scherer, 2010). For example, expressions of anger, fear, contempt, sadness, disgust and misunderstanding might all be associated with a frown and consequently recognised by the annotation scheme as ‘unpleasant emotions’. They are, however, cues to fundamentally
different emotional states and need to be dealt with differently by a health practitioner communicating with a patient. In other words, the Cues & Concerns approach only captures a tiny fraction of emotional states and attitudes that can be detected in doctor-patient interactions (Henry et al., 2011). Moreover, it only focuses on the patient’s behaviour and does not in any way address the nonverbal behaviour of the medical practitioner. However, the ability to correctly encode appropriate nonverbal cues is just as important as the ability to decode them (Knapp & Hall, 2010).

Given the issues raised above, the current techniques of medical interviewing analysis appear to be in dire need of a systematic approach to annotating nonverbal behaviour of both patients and health practitioners. Such an approach could offer a more comprehensive picture because it would take into consideration both cues to emotions, as well as cues to other signals of social communication and interpersonal regulation. These might serve as signals of competence and power (dominance), or of affiliation (Troisi, 1999; Vinciarelli et al., 2008), all of which play an important role in doctor-patient interactions (Ambady et al., 2002).

Genetic Counselling
In the first part of the project, we explored the use of (semi-)automatic annotation techniques for the analysis of video-recordings collected in the context of genetic counselling (Pieterse, 2005; Pieterse et al., 2007). For vocal analyses, a dataset consisting of 626 vocal fragments from 27 counselees and 14 counsellors was analyzed with Praat 5.3 (Boersma & Weenink, 2013). Since vocal pitch is an important source of information about speaker emotions (Juslin & Laukka, 2003), we extracted several pitch parameters (Mean, SD, Min, Max and Range), using the standard autocorrelation method (Boersma, 1993). Like sadness, expressions of anxiety/worry are typically presented with lower pitch (Goudbeek & Scherer, 2009) and in fact, in our dataset, lower pitch mean and max in counselees speech correlated with higher degrees of reported anxiety. We also found support for the view that for health care providers, listeners perceive lower pitch (and reduced rate) as more caring and sympathetic (McHenry et al., 2012): In the counsellor vocal profiles, higher, more varied pitch correlated negatively with reported satisfaction with the counseling visit. In general, the results show that automatic analyses of vocal cues can be used to analyze and, potentially, improve medical interactions. For facial expressions, we are currently evaluating the success of several algorithms (CERT, Face Tracker, and Frame Differencing Methods) designed to detect facial action units and/or basic emotions. Following the automatic analyses, their outcomes will be linked to existing RIAS annotations available for the dataset.

References


Social Signal Processing is the first book to cover all aspects of the modeling, automated detection, analysis, and synthesis of nonverbal behavior in human-human and human-machine interactions. Authoritative surveys address conceptual foundations, machine analysis and synthesis of social signal processing, and applications. Foundational topics include affect perception and interpersonal coordination in communication; later chapters cover technologies for automatic detection and understanding such as computational paralinguistics and facial expression analysis and for the generation of artificial Social signal processing, the detection and interpretation of temporal patterns of nonverbal behavioral cues in a given context, is a natural and, for the most part, unconscious process for humans. This process remains a rather difficult task for machines, but it is an important one to achieve if the goal is to realize a naturalistic and efficient human-computer interaction. Agreements and disagreements are social attitudes that occur daily and are inevitable in a variety of everyday situations, from everyday simple discussions over, e.g., which restaurant to dine in, to traditionally controversial. Evidence suggests that integrating a human supporter into such services mitigates these challenges, however, it remains understudied how supporter involvement benefits client outcomes, and how to maximize such effects. We present our analysis of 234,735 supporter messages to discover how different support strategies correlate with clinical outcomes. 2. Our work indicates that concrete, positive, and supportive messages from supporters that reference social behaviors are strongly associated with better outcomes; and that the importance of support strategies can vary dependent on a client’s specific context (e.g., their mental health, platform use). Spatially integrated social science is a broad term used to describe the integration of space and place in social science research using Geographic Information Systems (GIS). These processes include strategies that prioritize local relationships to ensure continuity of access to land for operational purposes [23]. Mining companies maintain these relationships primarily through preferential recruitment and inclusion in supply chains, social investment, and routine stakeholder engagement. Under the United Nations Guiding Principles on Business and Human Rights [28], for example, companies are required to exercise human rights due diligence in order to know and show that they respect human rights. Recently published articles from Biomedical Signal Processing and Control. A deep convolutional neural network for the detection of polyps in colonoscopy images. Tariq Rahim | Syed Ali Hassan | ... New affiliation with the European Alliance of Medical and Biological Engineering and Science. Biomedical Signal Processing and Control is now part of Elsevier’s Article Transfer Service (ATS)! View All. Below is a recent list of 2020–2021 articles that have had the most social media attention. The Plum Print next to each article shows the relative activity in each of these categories of metrics: Captures, Mentions, Social Media and Citations. Go here to learn more about PlumX Metrics. A convolutional neural network approach to detect congestive heart failure.