Accepted for publication at the AHFE 2010 conference

CHAPTER 1 [TIMES NEW ROMAN SIZE 16]

Enabling pre-hospital documentation via spoken language understanding on the modern battlefield

Lisa Anthony, Kenny Sharma, Susan Harkness Regli,
Kathleen Stibler, and Patrice D. Tremoulet
Lockheed Martin Advanced Technology Laboratories
3 Executive Campus, Cherry Hill, NJ, USA

LTC Robert T. Gerhardt
U.S. Army Institute of Surgical Research
3400 Rawley Chambers Avenue, Fort Sam Houston, TX, USA

ABSTRACT

Lockheed Martin Advanced Technology Laboratories (LM ATL) and the U.S. Army Institute of Surgical Research (USAISR) have been collaborating to improve methods of pre-hospital documentation. The collaboration is focused on creating a pre-hospital documentation grammar and spoken natural language understanding capability to support a hands-free and eyes-free interaction paradigm. This interaction paradigm will minimize the impact on the field medic during treatment of patients in stressful combat situations. The MediTRA-PH (Medical Treatment and Reporting Assistant: Pre-Hospital) prototype is a proof-of-concept of this approach. This paper presents the user-focused concept of operations for MediTRA-PH, the design and development cycles undertaken and preliminary evaluation results in terms of recognition accuracy and user satisfaction.

Keywords: Pre-hospital documentation, Military medical care, Trauma care, Field medics, User-centered design, Spoken language understanding, SLICE.

INTRODUCTION

Pre-hospital trauma care in the military theater is the first link in a chain of medical care as a casualty moves from point of injury to combat support hospitals and eventually to larger facilities. At each transfer point, there is a high risk of information breakdown. Beginning at the patient's point of injury, there is a significant lack of consistent documentation of medically-relevant information, including: mechanism of injury, demographics, medical interventions performed, and outcome. Field medics, lacking any other resource, often resort to writing the details of the patient's identity and treatments on bandages or medical tape on the patient's skin. Lockheed Martin Advanced Technology Laboratories (LM ATL) and the US Army Institute of Surgical Research (USAISR) have collaborated on a prototype system, aimed at improving pre-hospital documentation. This system, MediTRA-PH, or the Medical Treatment and Reporting Assistant: Pre-Hospital, uses a hands-free, eyes-free spoken-language interaction paradigm. This design choice was inspired by military medical subject matter experts (SMEs) input that, while early and frequent documentation does help medical providers later in the treatment chain to be more effective, field medics will only comply with documentation requirements that do not interfere with treatment and saving lives. MediTRA-PH uses LM ATL's Spoken-Language Interfaces for Computing Environment (SLICE) spoken-language understanding technology as the primary method of interaction with the system. This paper presents work to date on MediTRA-PH, including consultation with military medical SMEs on the spoken language interface to the system, and preliminary evaluation results: recognition accuracy and user satisfaction.

RELATED WORK

CURRENT MILITARY MEDICAL SYSTEMS

Since 2003, the US military has deployed a progressive series of components under the Medical Communications for Combat Casualty Care (MC4) system. Goals of the MC4 system include "enabling lifelong electronic medical records, streamlined medical logistics and enhanced situational awareness for Army tactical forces.[1]" MC4 consists of many different inter-operable systems and each focuses on a different aspect of the healthcare process. Handheld devices are carried by medical personnel, allowing them to enter information about initial treatment procedures so that it becomes part of the patient's electronic medical record. However, military medical personnel we consulted cited MC4's point-of-injury care support as

[1] https://www.mc4.army.mil/about.asp#how

cumbersome. Also, interacting with MC4 systems can interfere with patient interaction because doctors and other medical personnel are required to split their attention between the system and the patient. Both of these limitations are targets of the MediTRA-PH prototype system described in this paper. MediTRA-PH will be a hands-free, eyes-free, point-of-injury component of MC4 that integrates with the electronic medical records (EMR) pipeline created and enabled by MC4.

SPOKEN LANGUAGE UNDERSTANDING SYSTEMS

To build the MediTRA-PH prototype, we leverage our spoken language understanding capabilities, called Spoken Language Interaction for Computing Environments (SLICE), and our experience building multimodal systems for the Office of Naval Research (ONR) and DARPA. Based on nine years of spoken language understanding research begun under the DARPA Communicator program, SLICE (Daniels & Hastie, 2003) provides hands-free operation of computer systems, allowing warfighters in the field to use voice to interact with distributed information sources. Other spoken language systems have been developed for the medical domain, such as Q-Med (Johnson et al, 1992) and Dragon Medical (Nuance, 2008). However, these systems have focused on recognition rather than understanding and are used primarily for dictation and diagnostic interviews during office visits. The environments in which MediTRA-PH is used impose much more demanding challenges on the software, including ambient noise levels, faster pace, and potentially nonresponsive patients.

MEDITRA-PH PROTOTYPE

DESIGN PROCESS AND METHODOLOGY

We engaged in a hybrid user-centered design process called Interface Design for Engineering and Advanced Systems (IDEAS) (Regli & Tremoulet, 2007) to create a demonstration interface that shows the spoken-language understanding in action. IDEAS, an extension to traditional User-Centered Design (UCD), directly facilitates discussion and translation between engineering subject matter experts and domain subject matter experts. The LM ATL team worked with military medical personnel early and often while designing the MediTRA-PH prototype to ensure the eventual prototype system would meet target end-user goals. USAISR arranged a two-day workshop early in the effort to accomplish two primary goals: (1) to elicit the context in which the system would be used; and (2) to elicit specific grammar and vocabulary requirements. USAISR provided four field medics (two from the US Army, and two from the US Army Ranger Regiment), one US Army nurse, one US Army flight surgeon, and one US Army Reserve physician to consult on this panel.

Day one of the workshop was organized as a focus group, in which the LM ATL team asked directed questions designed to elicit domain context information and to engage the medical SMEs in the process of developing a vision for the MediTRA-PH system. Day two involved engaging the field medics in scenario role-playing activities to allow the system designers, as well as all stakeholders, to witness re-enactments of several trauma care incidents in as close to the natural context as possible. Two of the medics each performed two scenario re-enactments of actual events they experienced (minus any identifying details), in two situations: (1) care under fire and (2) tactical field care. These re-enactments revealed that a fielded MediTRA-PH system would have to be extremely robust, flexible, and ruggedized, as well as passive and non-intrusive, in order to successfully integrate with the field medic's environment and context, confirming the utility of a hands-free and eyes-free spoken-language interaction. To inform the design of this spoken-language interaction, we used both direct and indirect elicitation methods to define the vocabulary and grammar that medical personnel use in the field. In the structured interviews, we asked the field medics to directly report commands that they would like to give to the system. In the re-enactments, we recorded the language that they used in context. Both types of elicitation are important because users often cannot explicitly describe how they do a task once it becomes second nature (Beyer & Holtzblatt, 1998). After the workshop, the LM ATL team sent follow-up questions and intermediary design artifacts to the medical SMEs for review and comment, to keep them involved and engaged as the prototype was iteratively developed.

PROTOTYPE CONCEPT OF OPERATIONS

Based on the workshop, we developed the following concept of operations for the MediTRA-PH system. A warfighter sustains an injury during contact. As a field medic moves within treatment distance, the medic's ruggedized handheld platform buzzes to indicate connection with the patient's electronic information carrier (EIC), similar to a traditional identification tag, or "dog tag." The medic says "get PHR" (Personal Health Record), signaling his MediTRA-PH device to query the patient's digital dog tag for complete medical history and present a visual summary of the most important elements. The medic can also request that the information be read in his headset, or ask to be warned only if a proposed treatment is contraindicated based on prior medical history. As the medic treats the patient, he speaks aloud his treatment steps, and MediTRA-PH passively captures the information. It adds this information to the incident report as well as the patient's PHR. Throughout the treatment, the patient's digital dog tag receives updates from MediTRA-PH to ensure that documentation from this point of injury treatment care episode is carried forward when the patient is evacuated from the area. If the medic is working on several patients, as in for example a mass casualty event, the digital dog tag syncing solution ensures documentation is mapped to the right patient. When the medic returns to base, he docks his handheld platform to a MediTRA-PH-enabled machine, which can assist him in filling out his required paperwork, leveraging the data captured from the field. Using the desktop interface, field medics can update MediTRA-PH's vocabulary as new treatments and medications are deployed.

RECOMMENDATIONS FOR REPORTING SYSTEMS

Although the main product of this effort is a prototype system, another important result is the generation of design recommendations for medical domain field reporting systems. These design recommendations were developed as a result of our interactions with medical SMEs during this effort, and reflect the participatory, collaborative nature of our design process, which involves end users and technologists in developing a user-centered vision and, ultimately, system. We have divided the recommendations into three areas: (1) general interaction preferences; (2) spoken language reporting; (3) other desired features and functionality.

General Interaction Preferences

The medical personnel strongly recommended a system designed to seamlessly integrate into existing treatment episodes via a passive hands-free and eyes-free interaction paradigm. Medics use both their hands to treat patients. A trauma care reporting system should be usable without physical interaction. Furthermore, medical personnel want to maximize the quality of the time they spend with their patients and are unwilling to split attention between a computer and a patient, or ignore a patient completely for a minute or two as a report is filled out. Because of the speed and urgency of treatment episodes during pre-hospital trauma care, a system like MediTRA-PH must be configurable to provide feedback on-demand or not at all. Entering information into the PHR must also be extremely rapid so as not to interfere with treatment. Point-of-care documentation systems must also support fluid transition between multiple patients for one or more medics. In mass casualty events, a medic might be treating multiple patients interchangeably, jumping back and forth quickly between injuries. The system must be enable rapid transition, preferably automatically if possible, so that information about injuries and treatments is matched to the correct patient in the electronic medical record (EMR). Finally, point-of-care documentation systems must support interactions for both experienced and novice medics. According to our data from the workshop, 6500 new medics are trained every year at Ft. Sam Houston, leading to a large influx of deployed medical personnel who may be seeing battle injuries for the first time. Stress and inexperience sometimes lead medics to forget important elements of their training. A system like MediTRA-PH should support treatment guides and reminders or alerts when the medic requests it or the system detects him faltering. Rapid deployment is crucial to the military medical provider, so a system like MediTRA-PH should require as little end-user training as possible.

Spoken-Language Reporting

The key reporting modality that can support all interaction preferences and fit well into the context of trauma care is spoken-language reporting. As our military medical SMEs told us, many medics already speak aloud while treating patients, and most recognize that this is beneficial in order to keep focused on the moment and also to interact with the patient to keep him conscious. To best support the medic and require the least amount of user-training, a medical voice documentation system must focus on keywords people use in the field already, such as treatment devices and medication names, as a means of interaction with the system, rather than introducing new language. Because medics vary in experience, it is important for a system like MediTRA-PH to support multiple ways to report a similar thing, including using expert medical terminology (i.e., “anterior femur”) and layman’s terms (i.e., “right thigh”).– Also, medical terminology changes frequently as new treatment devices and medications are deployed, and a system for documenting pre-hospital trauma care must be able to handle all current vocabulary. Furthermore, individual medics might use shorthand or other localized terminology that vary from user to user. An essential design recommendation is to provide an end-user focused interface for vocabulary editing and refining that can be used by medics to add, remove or change vocabulary terms and potentially even new reporting structures to the extent possible with current technological capability. Finally, the spoken-language understanding mechanism must be able to handle a wide variety of speakers, including people with accents, male or female voices, and voices of people in situations of stress, which could change their speaker characteristics. Furthermore, it must be able to listen passively and only capture language that is medically relevant, ignoring and not recording off-topic conversations with other medical staff or the patient, for reasons of privacy.

Other Desired Features and Functionality

Although the primary context of use for a system like MediTRA-PH is the point of injury, patient care documentation does not end when the medic leaves the battlefield. Medics need to be able to enter and edit reports when they have returned to the base, to ensure that all information was captured accurately and completely. Medics have paperwork obligations that a system like MediTRA-PH can make less tedious by feeding directly into the medic’s incident report forms such as the SF-600 from the spoken-language report. A system like MediTRA-PH must integrate into the existing electronic medical records pipeline (i.e., MC4), supplying information about trauma care into a patient’s long-term health record.

DEVELOPMENT

A working prototype was built for demonstration and evaluation of the MediTRA-PH approach, based on LM ATL’s SLICE architecture. SLICE converts spoken audio into a representation of the context-dependent meaning, or user intent. The underlying language model representation provides a rule-based classification system that enables the system to process spoken language input to perform tasks.

Architecture and Language Model Approach

The MediTRA-PH SLICE implementation uses an XML-based language model that defines the medic domain in terms of vocabulary and contextual relationships between categories of concepts. Figure 1 shows a diagram of how the language model concepts are related to each other.

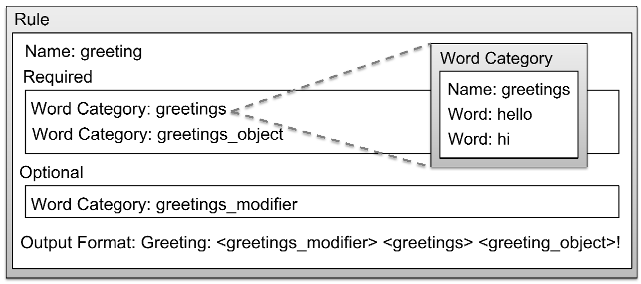

Figure 1. Language model used in the MediTRA-PH prototype system.

Words are the base unit of the language model. A word consists of a literal text representation and one or more pronunciations. Word Categories are logical groupings of particular types of words that can be used interchangeably. The words may have different meanings, but represent the same general concept in relation to each other. For example “fractured” and “broken” can both be categorized as breaks in bones. Rules define logical groupings of categories that must be spoken together to have meaning. Each rule provides both optional and required categories, allowing for flexibility in how much information is required to satisfy a particular rule. The order in which words in each category are spoken is unimportant, compensating for natural variation in how individuals describe the same event or object. Rule Mappings define how to route the data within the SLICE architecture. Particular rules can be handled differently than other rules; for example, an injury should be routed and handled differently than a system command.

Live Speech Recognition and Segmentation

For the initial prototype development, it is assumed that the field medic user will use a key phrase (several variations on “start / stop recorder”) to begin and end a particular patient’s record. Once the key phrase is uttered, MediTRA-PH does not require any further explicit segmentation of words, phrase or sentences. In order to support the kinds of language our military medical SMEs delivered, the speech recognizer is configured to interpret both continuous words and bursts of phrases, rather than requiring a complete phrase before trying to recognize. This configuration results in a constant streaming of text recognized from the user’s raw audio, which is matched against one underlying language model. MediTRA-PH’s language model is constructed to minimize the amount of information required for each phrase to be successfully understood, while also minimizing potential conflicts (in the form of redundancy and overlaps) between phrases. Minimizing the complexity of language model components also attempts to reduce the amount of information lost through misrecognition or other errors. To identify the correct break between phrases, disambiguation and conflict management are performed, allowing for delineation of phrases at the earliest possible and logical point in order to prevent error propagation and maximize language model matching efficiency.

Proof-of-concept Graphical User Interface (GUI)

The MediTRA-PH prototype includes a proof-of-concept graphical user interface (GUI) to demonstrate possible visualizations of treatment episodes. Figure 2 shows four possible representations of information classified as injuries, treatments, and mechanisms of injury. In Figure 2, Section 1 illustrates a categorized display of information, representing the basic information types captured in separate fields for ease of readability. Section 2 illustrates a moving timeline of categorized information, where the update interval is user-configurable, which illustrates what was done and how long ago. Section 3 shows a standard body injury diagram that can be used to quickly assess where injuries are located. Section 4 shows a time-tagged transcript of all available information, which is valuable during the post-treatment analysis period. These display visualizations are only preliminary and require iterative user-oriented refinement to ensure they meet user needs.

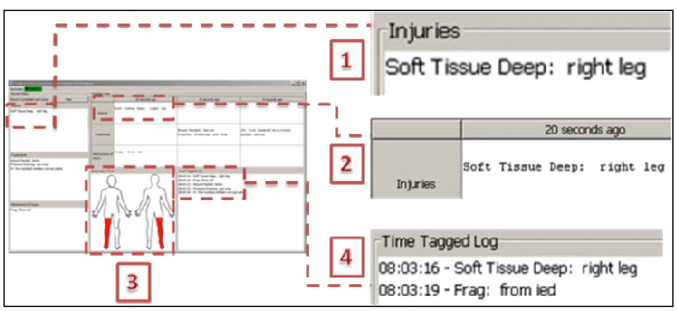

Figure 2. Proof-of-concept graphical user interface (GUI) for MediTRA-PH.

DEMONSTRATION AND QUALITATIVE FEEDBACK

The MediTRA-PH prototype was demonstrated during the 2009 Advanced Technology Applications for Combat Casualty Care (ATACCC) conference. Several formal sessions and informal discussions occurred with field medics of varying experience, including some of the SMEs consulted during initial concept development. Reaction to MediTRA-PH was positive. Specifically, the hands-free, eyes-free collection of speech was viewed as a viable method of nonintrusive information capture. We received positive confirmation that our supported language model was accurate and of high utility, and received constructive feedback on both additional vocabulary to support and information to display in the GUI. Overall, the SMEs consulted at ATACC indicated that use of the MediTRA-PH prototype was intuitive and met their conceptual expectations.

SPOKEN LANGUAGE UNDERSTANDING PERFORMANCE

To evaluate the performance of MediTRA-PH in terms of how well it can understand spoken medical language, we collected a corpus of in-domain spoken language from 15 participants with varying levels of medical domain familiarity. The corpus was comprised of individual treatment episodes with a total of 32 domain-relevant utterances. The audio files for each vignette and participant were fed into MediTRA-PH automatically and the results of the recognition step and the language model matching step were recorded. Finally, the results were compared to ground truth. Three possible outcomes for each utterance within a vignette were possible: (1) true positive, or the correct rule fired; (2) false positive, or an incorrect rule fired; and (3) false negative, or a failure to fire any rule. Because the corpus did not contain any out-of-domain utterances, there were no true negatives. The results of this analysis are as follows. Over all utterances in the corpus, 84.89% of them were correctly classified by MediTRA-PH, meaning the correct rule fired, enabling the system to automatically populate the patient's PHR from spoken language input. Of the remainder, only 0.68% were false positives, while most of the errors MediTRA-PH made were in the form of false negatives (14.43% overall), meaning that no rule was fired. This occurred because of the low spoken-language recognition accuracy achieved by MediTRA-PH's SLICE core in the medical domain; we computed the standard NIST metric, word error rate (WER) (Pallett et al, 1990), to be 73.93%. Speech recognition can be tuned to perform better in domain-specific applications if the recognizer models are trained to specific dictionaries. It is possible that by decreasing the SLICE word error rate, performance of MediTRA-PH at matching rules would significantly improve.

CONCLUSIONS AND FUTURE WORK

This paper has presented MediTRA-PH, a prototype enabling hands-free, eyes-free spoken language reporting of pre-hospital care. This prototype is the product of a multi-disciplinary collaboration between technologists, human factors specialists, and medical SMEs. If deployed, such a system would greatly alleviate the problem of loss of medical treatment information from point of injury contexts and improve patient treatment and hand-off. Moreover, the technology described here could be applied to other domains such as disaster relief, civilian health records, Homeland Security, and so on. Further work in this area could include the following: (1) refining the pre-hospital grammar through further work with medical SMEs; (2)

extending the SLICE speech recognition technology to perform better for domain-specific medical terminology, yielding an increase in recognition accuracy and therefore language model matching accuracy; (3) more formal evaluations of the refined MediTRA-PH prototype in lab and field exercises, including testing of out-of-domain vocabulary; (4) exploring user interface designs and visualizations that allow practical use of MediTRA-PH during conflict scenarios; and (5) expanding the scope of the work by consulting with other medical SMEs to validate our design recommendations with the broader military medical community.

ACKNOWLEDGEMENTS

The authors thank the military medical SMEs who participated in the workshops and demonstrations as part of this effort. Privacy concerns prevent us from listing them by name. This work was funded and carried out jointly by Lockheed Martin Advanced Technology Laboratories and the U.S. Army Institute of Surgical Research (USAISR). COL David G. Gilbertson, formerly of the USAISR, provided invaluable support to this research. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

REFERENCES

Beyer, H. and Holtzblatt, K. (1998), *Contextual Design*. Morgan Kaufmann Publishers, New York.

Daniels, J.J. and Hastie, H.W. (2003), “The pragmatics of taking a spoken language system out of the laboratory.” In *Proceedings of the HLT-NAACL 2003 Workshop on Research Directions in Dialogue Processing*, Volume 7, Human Language Technology Conference. Association for Computational Linguistics, Morristown, NJ, 16-18.

Johnson, K., Poon, A., Shiffman, S., Lin, R., and Fagan, L.M. (1992), *Q-Med: a spoken-language system to conduct medical interviews*. Knowledge Systems, AI Laboratory (KSL-92-09).

Nuance. (2008), *Optimizing clinical productivity: using speech recognition with medical features vs. a general-purpose solution*. Whitepaper: http://www.nuance.com/healthcare/pdf/wp_healthcare_MDComparative.pdf

Pallet, D.S., Fisher W.M., and Fiscus, J.G. (1990), “Tools for the Analysis of Benchmark Speech Recognition Tests.” In *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume 1, 97-100.

Regli, S.H. and Tremoulet, P.D. (2007), “The IDEAS Process: Interaction Design and Engineering for Advanced Systems.” *APA Division 21, Division 19, & HFES Potomac Chapter Annual Symposium on Applied Research*. George Mason University, Fairfax, VA, March 1-2.