Summarising Complex ICU Data in Natural Language: demonstration of the BT-45 system

François Portet, PhD¹, Albert Gatt, PhD¹, Ehud Reiter, PhD¹, Jim Hunter, PhD¹, Somayajulu Sripada, PhD¹, Cindy Sykes², Yvonne Freer, PhD², Neil McIntosh DSc², Marian van der Meulen, PhD³ and Robert Logie, PhD³.

¹ Department of Computing Science, University of Aberdeen, UK,
² Simpson Centre for Reproductive Health, Edinburgh Royal Infirmary
and ³ Department of Psychology, University of Edinburgh

Abstract

As ICUs generate increasing amounts of information, writing medical reports involves complex time-consuming reasoning to build a coherent text which will be meaningful to those who will use it for decisions making (e.g.: for nurse handover). Moreover, it has been shown that summarizing complex multi-channel physiological and discrete data in natural language (text) can lead to better decision-making in the intensive care unit (ICU). To facilitate this summarisation, as part of the BabyTalk project, we developed a system called BT-45 that automatically generates textual summaries from periods of continuous and discrete data in a neonatal ICU. The demonstration will show the system running on real data and will detail the steps in the construction of the final text. Although these summaries are not yet as good as those generated by human experts, we have demonstrated experimentally that they lead to as good decision-making as can be achieved through presenting the same data graphically.

Introduction

The system presented is related to the BabyTalk project¹ whose aim is to support clinical staff in a neonatal ICU by providing a textual summary of the state of a baby over a period of time; and to facilitate the writing of medical reports. The BT-45 prototype is an extension of the NLG pipeline architecture originally proposed in.² It automatically summarizes 45 minutes of continuous and discrete data in four stages. In an intensive care unit (ICU), the data available for a patient typically consists of: (i) continuously monitored physiological signals (such as heart rate) sampled every few seconds and (ii) discrete events (such as equipment settings, results of blood and other laboratory analyses). Physiological signals are processed by a Signal Analysis module to abstract the main features of the signals. A Data Interpretation module infers relations (causality, association, etc.) between signals features and other discrete events (lab results, medication, etc.) and abstracts patterns of events into higher description levels. A Content Determination module then selects the most important events and features, and aggregates them into a tree of linked events. Finally, the Micro Planning module translates this tree into coherent text. The modules are linked using an Ontology for the NICU of concepts acquired from clinicians. Few NLG systems consist of such a complex processing chain that goes from the signals to meaningful text.

BT-45 has been tested in an off-ward trial. Over thirty nurses and doctors were asked (individually) to say what clinical actions(s) they would take for a baby whose recent history over a period of about 45 minutes was presented either graphically (G), as text generated by a human expert (H) or as text generated by BT-45 (C). The proportion of appropriate actions minus the proportion of inappropriate actions chosen for each scenario served as a score for evaluating each mode of presentation. Results showed 0.33 (sd = 0.14) for G condition, 0.39 (sd = 0.11) for H condition and 0.34 (sd = 0.14) for C condition. H leads to significantly better decisions than with G and our computer system C performed at least as well as G (no significant difference). After only one year of development, this result encourages us to develop extended systems, specifically tailored to doctors, nurses and families of the newborn.