BabyTalk: A Core Architecture to Summarise ICU Data as Tailored Text

François PORTET\textsuperscript{a,1}, Albert GATT\textsuperscript{a}, Jim HUNTER\textsuperscript{a}, Ehud REITER\textsuperscript{a}, Somayajulu SRIPADA\textsuperscript{a}, and Feng GAO\textsuperscript{a}

\textsuperscript{a}Department of computing science, University of Aberdeen, ABERDEEN, Scotland, UK

\textbf{Keywords:} Systems Architecture, Knowledge-Based Systems, Decision Support, Patient Monitoring, Linguistics, Paediatrics, Primary Care

In Intensive Care Units (ICU), report writing is a routine clinical time-consuming task that involves complex reasoning and can be prone to errors. Furthermore, medical data is generally represented as graphs but it has been showed that, in certain circumstances, textual representations of data can lead to better decision making. The system we are developing aims at generating texts summarising period of baby’s continuous and discrete data in a Neonatal ICU (NICU). This is part of the BabyTalk project [1] which is a join effort of the University of Aberdeen, the Royal Infirmary of Edinburgh and the University of Edinburgh.

We have already developed a prototype [2] that has been tested in a clinical 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 our prototype (C). The result showed the value of C and H texts for decision support in comparison with G presentation.

We are now extending this prototype to generate texts for three different types of audiences: physicians (for decision support), nurses (for reporting and patient follow up) and the baby’s family (for communication support).

The poster presents the architecture of the extended system. The system consists of the following stages: data acquisition - extracting information from different databases into the knowledge base supported by a NICU ontology; data analysis - signal and text processing modules extracting information from the unstructured data; data interpretation - temporal reasoning, procedure recognition, etc. to relate the events together and abstract high level events (e.g. procedure recognition); content selection - selection and organisation of the relevant information according to the final user (nurse, doctor or family); text generation - translation into comprehensive text; communication - report, email, SMS.
