line of thought is steared by the statements, guidelines, and questions from the creativity techniques. Each generated idea (as for example in Figure 1) can be connected to another by an association with a label taken from the set of 5WH1 questions (labels on edges in Figure 1 for example). This can be modeled as graph. Further, each idea represented as a node can be supported by a sketch drawn in a sketch editor. Once a group is finished with the idea graph, a consolidation session might be called upon where the graph transformation of existing graphs, including the one generated in the latest session, are employed to achieve a final agreed upon suggestion for a solution.

The benefit of such an approach is that teams get user-friendly support for consolidation and preservation of the creativity session results. The results are far more than just snapshots of white boards; they can be used for later processing, such as automatic discovery of hidden implicit relations between ideas, recommendations of related existing ideas, guiding through an information space of ideas and so on. As a consequence more people are able to contribute ideas, to manage creativity and its results more transparently and efficiently, and to benefit from previous knowledge constructed in other creative sessions.

Link:
http://www.idspace-project.org

Please contact:
Peter Dolog
Aalborg University, Denmark
E-mail: dolog@cs.aau.dk

Extracting Information from Free-text Mammography Reports

by Andrea Esuli, Diego Marcheggiani and Fabrizio Sebastiani

Researchers from ISTI-CNR, Pisa, aim to effectively and efficiently extract information from free-text mammography reports, as a step towards the automatic transformation of unstructured medical documentation into structured data.

Information Extraction is the discipline concerned with the extraction of natural language expressions from free text, where these expressions instantiate concepts of interest in a given domain. For instance, given a corpus of job announcements, one may want to extract from each announcement the natural language expressions that describe the nature of the job, annual salary, job location, etc. Put another way, information extraction may be seen as the activity of populating a structured information repository (such as a relational database, where "job", "annual salary" and "job location" count as attributes) from an unstructured information source such as a corpus of free text. Another example of information extraction is searching free text for named entities, ie, names of persons, locations, geopolitical organizations, and the like.

An application of great interest is extracting information from free-text medical reports, such as radiology reports. These reports are unstructured in nature, since they are written in free text by medical personnel. However, applying

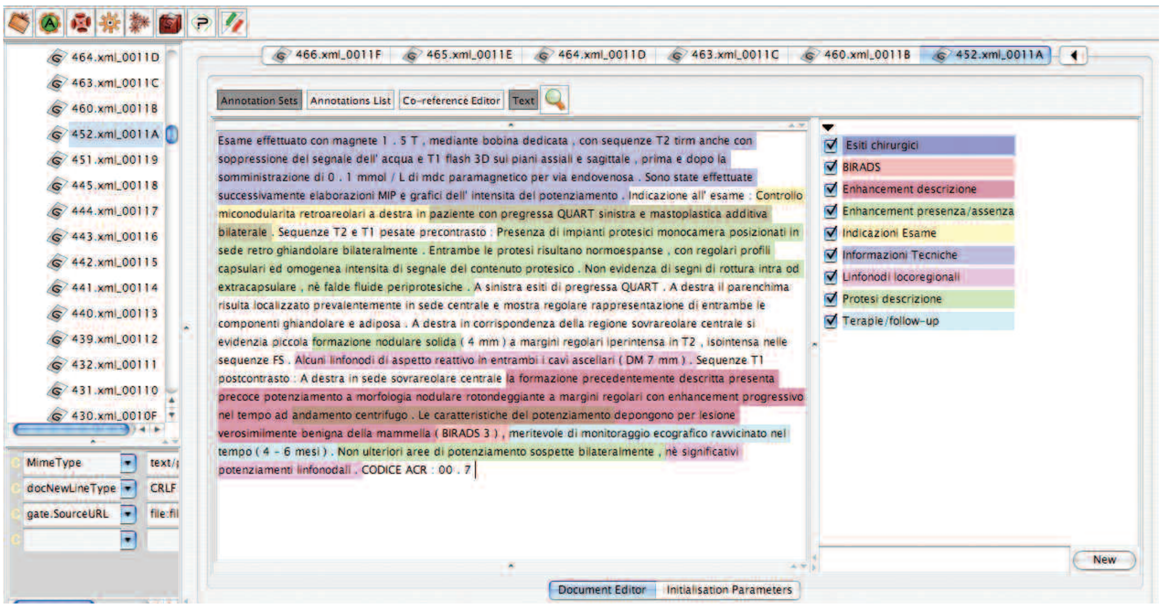

Figure 1: A mammographic report automatically annotated according to the nine concepts of interest.

information extraction would be beneficial, since extracting key data and converting them into a structured (eg, tabular) format would greatly contribute towards endowing patients with electronic medical records that, aside from improving interoperability among different medical information systems, could be used in a variety of applications, from a patient's care to epidemiology and clinical studies.

In recent months we have worked on a system for automatically extracting information from mammography reports written in Italian. There are two main approaches to designing an information extraction system. One is the rule-based approach, which consists of writing a set of rules which relate natural language patterns with the concepts to be extracted from the text. This approach, while potentially effective, is too costly, since it requires a lot of human power to write the rules, which are jointly written by a domain expert - say, an expert radiologist - and a natural language engineer. We have followed an alternative approach, which is based on machine learning. According to this approach, a general-purpose learning software learns to relate natural language patterns with the concepts to be instantiated from a set of manually annotated free texts, ie texts in which the instances of the concepts of interest have been marked by a domain expert. The advantage of this approach is that the human power required to annotate the texts needed to train the system is much smaller than that needed to manually write the extraction rules.

The system we have built uses "conditional random fields" (CRFs) as a learning technique. CRFs were explicitly devised for managing data of a sequential nature, such as text, and have given good results on text-related tasks such as named entity extraction and part-of-speech tagging.

We have tested our system on a set of 405 mammographical reports written (in Italian) by medical personnel of the Radiology Institute of Policlinico Umberto I of Rome, and manually annotated by two expert radiologists of the same institution according to nine concepts of interest (eg, "Followup Therapies"). 336 reports were annotated by one radiologist only, while the other 59 were independently annotated by both radiologists. The presence of reports annotated by both radiologists has allowed us to directly compare the accuracy of our system with human accuracy, by comparing the agreement between the system's and the radiologist's annotations with the agreement between the annotations of the two radiologists. Our experiments, run by 10-fold cross validation, have shown that our system obtains near-human performance: the agreement between system and human, measured by the standard "macroaveraged F1" measure, turned out to be 0.776, while the agreement between the two experts was 0.794 (higher values are better, since 0 and 1 indicate perfect disagreement and perfect agreement, respectively). These results are especially encouraging because no specialized lexical resource was used in the experiments, since no such resource exists for the radiological / mammographic sector for Italian.

Please contact:
Fabrizio Sebastiani
ISTI-CNR, Italy
E-mail: fabrizio.sebastiani@isti.cnr.it

Parental Control for Mobile Devices

by Gabriele Costa, Francesco la Torre, Fabio Martinelli and Paolo Mori

As a result of the rapid increase in mobile device capabilities, we believe that many users will soon migrate most of their tasks from the computer to the smart phone. We thus feel that new, more effective security mechanisms are needed to protect users, especially minors. For this reason, we are currently engineering a software suite designed to monitor all security-relevant activities and, where necessary, block unsuitable material.

In western countries mobile phones now outnumber people. The majority of these phones have embedded cameras and good connectivity support, connecting to the Internet via their mobile operator network or in wireless mode, and communicating directly with other devices through a Bluetooth interface. This means that the gap in functionality between personal computers and smart phones is being continuously reduced. Mobile devices such as smart phones or Personal Digital Assistants (PDAs) are now powerful enough to run applications that are comparable to those running on normal PCs.

The mobile phone is thus now typically employed for many tasks, such as browsing the Internet, reading and writing e-mails, exchanging data (photos, contacts, files, etc) with other devices, and so on. Clearly, these operations bring with them new vulnerabilities and raise new threats to the user's privacy. A particular source of concern is the widespread use of mobile devices by young people and minors who can easily access content that is inappropriate for their age. Filters provided by the mobile telco operators are not sufficient to block such material because they only inspect data sent via the telco network and have no control on direct connections (eg Bluetooth). As a consequence, in recent years, much research has focused on the provision of customised, security support for mobile devices. However, what is still missing is a general security framework that controls applications as well as phone calls and message content.

In order to meet this need, by calling upon the experience gained in the EU project S3MS (Security of Software and Services for Mobile Systems), we have developed a new modular security monitor support for mobile devices which uses a plugin-based, extensible architecture to meet the demands of security requirements which change from device to device and from user to user.

Our system aims to monitor all security-relevant activities involving a mobile device. The basic idea is to develop a minimal security core that evaluates the platform security state, accepts customised security policies, and deals with heterogeneous system events.

The system can be easily extended with new modules. When, following authentication, a Module M is plugged into the system. It injects the list of its security-relevant events and