[Jantuan, 2007] E.Jantuan. Knowledge validation tools. Bit+. The 7th International Conference "Information Technologies 2007", 24 -28 April, 2007, Chisinau, Republic of Moldova (to appear).

[Dzung, 1998] Palm Dzung L., Xu Chenyang, Prince Jerry L. A survey of current methods in medical image segmentation // Technical report. – Johns Hopkins University, Baltimore, 1998. http://www.csem.duke.edu/seminars/segmenation.pdf

[Jardim, 2002] Jardim Sandra Vilas Boas, Figueiredo Mário A. T. Automatic analysis of fetal echographic images // Proc. Portuguese Conf. on Pattern Recognition. – Aveiro: RecPad, 2002.

[Dinggang, 2003] Dinggang Shen, Yiqiang Zhan, Davatzikos C. Segmentation of prostate boundaries from ultrasound images using statistical shape model // IEEE Transactions on Image Processing. – 2003. – vol. 22, № 4. – P. 539-551.

[Popcova, 2004] O.Popcova, S.Cojocaru, C.Gaindric. Image Processing in Ultrasound Diagnostic System. 8th International Symposium on Automatic Control and Computer Science, Iasi, Romania, October 22-23, 2004.

Authors' Information

Svetlana Cojocaru *– Dr.hab., Institute of Mathematics and Computer Science, Academy of Sciences of Moldova, Academiei str., 5, Chisinau, Republic of Moldova, e-mail: sveta@math.md*

Constantin Gaindric *- Dr.hab., Prof.,corr.member Academy of Sciences of Moldova, Institute of Mathematics and Computer Science, Academy of Sciences of Moldova, Academiei str., 5, Chisinau, Republic of Moldova, e-mail: gaindric@math.md*

AUTOMATED RESPONSE TO QUERY SYSTEM

Vladimir Lovitskii, Michael Thrasher, David Traynor

Abstract: *SMS (Short Message Service) is now a hugely popular and a very powerful business communication technology for mobile phones. In order to respond correctly to a free form factual question given a large collection of texts, one needs to understand the question at a level that allows determining some of constraints the question imposes on a possible answer. These constraints may include a semantic classification of the sought after answer and may even suggest using different strategies when looking for and verifying a candidate answer. In this paper we focus on various attempts to overcome the major contradiction: the technical limitations of the SMS standard, and the huge number of found information for a possible answer.*

Keywords: *mobile text messages, text message analysis and question-answering system*

ACM Classification Keywords: *I.2 Artificial intelligence: I.2.7 Natural Language Processing: Text analysis.*

Introduction

This paper represents results of our further research in the text data mining and the natural language processing areas [1-5] restricted by mobile's text-based SMS messaging. SMS is now a very powerful business communication technology widely used from small businesses and home users through to large corporations, governmental and non-governmental organisations. However, many of these users have little or no experience of SMS technology and only a vague idea of how successful they could be when properly harnessing the power of SMS communication.

SMS text messaging is currently being evaluated in many different areas:

- **Mobile Banking.** Mobile Banking Services including: Account Balance Enquiry, Account Statement Enquiries, Cheque Status Enquiry, Cheque Book Requests, Fund Transfer between Accounts Credit/Debit Alerts, Minimum Balance Alerts, Bill Payment Alerts, Bill Payment, Recent Transaction History Requests, Information Requests i.e. Interest Rates/Exchange Rates. (e.g. the HSBC's SMS Enquiry Service [6]). Although these services are appearing they do raise specific issues concerning security, especially when the data is extremely sensitive and confidential. 2ergo have recognised this and have developed and patented a secure messaging protocol [7] to address such issues through encryption technology.

- Mobile Government services [8]. For example, Ireland's tax collection agency, Office of the Revenue Commissioners, as of 2005 now receives at least as many enquiries by text message as by telephone. The SMS enquiry service allows citizens to claim tax credits and request a number of tax forms and information leaflets by sending text messages from their mobile phones.
- Mobile Learning (SMS in education [9]). Students find SMS messages useful and are keen to use it in a number of ways: announcements, assessment marks, assessment feedback, appointments, revision tips and also quick easy access to some library service. Thus, it is believed that SMS in education can facilitate collaboration, strengthen community spirit, assist timely completion of coursework and other assessments, and ultimately reduce the attrition rate amongst students. 2ergo have released a new product targeted at the education market [7]. It is aimed at anti-truancy, anti-bullying and group messaging to provide school administrators with the means to manage and control these very important issues.
- Mobile Airport services [10]. The SMS service gives passengers fast responses to queries about the airport when they are possibly in transit and scheduled to connect with a flight. It reduces the need for people to phone the airport information number, which can often be engaged and thereby increasing anxiety levels amongst passengers, or go to the information desk inside the terminal.

Given the recent experience the list of areas where mobile messaging will prove useful is likely to be extended. What these applications have in common is that they represent a Well-Defined Application Domain (WDAD). For each of them a 5-digit short code can be predefined, e.g. when the message is sent to the 64222 short code WDAD "Edinburgh Airport" will be selected automatically and therefore the answer to the question *"Where is the nearest hotel?"* will be synthesized relatively simply by the Question-Answering System (QAS).

The central question to be addressed by this paper, however, is how to provide the response to a question if the application domain is unknown i.e. Unknown Application Domain (UAD). The attempt to find a response for UAD has been undertaken already by Google [11]. The goal of the Google SMS service is to provide the large existing base of users with access to the types of information they are most likely to need when mobile. Users simply send their query as a text message and receive results in the reply. Simple conventions are used to express queries for phone book listings, dictionary definitions, product prices, etc. Google touts this service as ***"Just text. No links. No web pages"*** [10]. Simply the answers one is looking to find. Examples of queries are *"pizza 21228"* to find pizza places located near to the University of Maryland, Baltimore County (UMBC), or *"george bush, washington dc"* to find the address and phone number of the US president or *"Price ipod 20gb"* to get a list of prices (and sellers) for an ipod. The Mobile Query (MQ) is sent to the 5-digit US short code 46645, which corresponds to Google on most phones. One could, of course, leverage existing mobile technology through the WAP browser on mobile phones as an alternative to using SMS. For example, if you want to find out the dictionary definition of the word *"spring"* simply enter message ***"define:spring"*** to search in your mobile "Google" on WAP enabled phones. The following response will be displayed: *"the season of growth; "the emerging buds were a sure sign of spring"; "he will hold office until the spring of next year""*. If the same query is repeated but on a personal computer rather more information regarding the definition will be displayed (the most compact part of those information is shown below).

Definitions of spring on the Web:

- the season of growth; "the emerging buds were a sure sign of spring"; "he will hold office until the spring of next year";
- a natural flow of ground water;
- jump: move forward by leaps and bounds; "The horse bounded across the meadow"; "The child leapt across the puddle"; "Can you jump over the fence?";
- form: develop into a distinctive entity; "our plans began to take shape";
- a metal elastic device that returns to its shape or position when pushed or pulled or pressed; "the spring was broken";

- leap: a light, self-propelled movement upwards or forwards;
- bounce: spring back; spring away from an impact; "The rubber ball bounced"; "These particles do not resile but they unite after they collide";
- give: the elasticity of something that can be stretched and returns to its original length;
- develop suddenly; "The tire sprang a leak";
- a point at which water issues forth;
- produce or disclose suddenly or unexpectedly; "He sprang these news on me just as I was leaving".

Despite the quite impressive results using MQ there remain some significant problems:

- Quality not quantity. The best conformity between the returned response and the MQ is more important than the quantity of information found. How can we define the quality of response? When a user asks for the definition *of spring* Google sent the first line of those definitions to the user's mobile. Is it what the user anticipated, or would it be more important for him to know that *"spring is a metal elastic device…"*? Using a single SMS restricts the answer to 160 characters, and depending upon the type of phone used the display area is very likely to be much less than this. Although *"no answer is worse than a bad answer"*. A disappointing response could deter the customer from using the SMS service again.
- No dialogue. Let us distinguish between QAS and Mobile QAS (MQAS). The principal difference is that QAS allows for a clarification dialogue [3,4] with the user to provide the best possible response through interaction but MQAS does not. As for MQAS only a single short user MQ must be used to create the response.
- Ungrammatical MQ. The dream about MQ: *"Be precise and informative about your problem. Write in clear, grammatical, correctly-spelt language"* – is far from being realised. As a rule an MQ will be ungrammatical, not because users are illiterate but because most users are lazy i.e. they want to achieve the desired result by using the minimum effort. For example, they do not want to use upper case to type MQ *"george bush, washington dc"* [9] or use dots to separate *"d"* and *"c"*. Other examples of typical grammatical errors are: typing *"its"* instead of *"it's"*, *"lose"* instead of *"loose"*, or *"discrete"* instead of *"discreet"*. The fact is that *MQ simply cannot be spelt, punctuated, and capitalised correctly* but the main requirement for MQAS is - to handle non-standard or poorly formed/structured (but, nevertheless, meaningful) user's MQ.
- Fuzzy MQ. The users are generally unable to describe completely and unambiguously what it is they are looking for [4]. For example, the MQ *"Who won the 2006 World Cup?"* is absolutely clear from the user's point of view because it goes without saying that *football* is implicit in the question. Fuzzy MQ requires fuzzy searching. Fuzzy response searching is a technique for finding a reply that approximately matches the search MQ.
- Facts Ambiguity. The problems of answering the MQ depends not only on an incomplete and/or ambiguous MQ but also on *facts ambiguity*, especially when MQAS, as a result of searching information on the internet, expands its local KB by adding appropriate facts. It is not easy to recognise *semantic*, *lexical* or *structural* ambiguity of facts. For example, the fact *"Stolen painting found by tree"* has structural ambiguity and can be interpreted as: (1) *A tree found a stolen painting*; (2) *A person found a stolen painting near a tree*. The MQ *"Who stole a painting?"* requires some effort from MQAS to reply correctly (we will discuss this in the *Mobile Queries Classification* session).

The meaning of an MQ depends not only on the things it describes, explicitly and implicitly, but also on both aspects of its causality: what caused it to be said and what result is intended by saying it. In other words, the meaning of an MQ depends not only on the MQ itself, but on who (age, sex, nationality) sends it and when, where, why and to whom it is sent. Such information helps to understand *what* the user meant in his/her MQ. In this paper we shall discuss why the user sent the MQ and the determination of some of the constraints the MQ imposes on a possible answer.

Why was the Mobile Query Sent?

Why do people not ask questions? It is very important to know why, especially if you want to learn and improve performance in an organisation. There are, as minimum, some 10 reasons why people never ask questions [12]. But here we want to find out the answer to the opposite question: *why do people ask questions?* While it is possible to question without thinking, it is impossible to think without questioning. A good question might just provide the means for overcoming a particular obstacle and achieving a stated goal or objective. There are some 12 reasons identified for asking questions [13]. This list is based on three assumptions: (1) a genuine question is one to which the answer is not known; (2) curiosity is the driver; and (3) the questions are intended in a constructive manner, i.e. to intimidate, dominate, demonstrate how smart one is, or to prove that one is right. These reasons are:

1. Gather information.
2. Build and maintain relationships, participate in communication.
3. Learn, teach, and reflect.
4. Think clearly, critically, and strategically.
5. Challenge assumptions.
6. Solve problems and make decisions.
7. Clarify and confirm information heard.
8. Negotiate and resolve conflicts.
9. Set and accomplish goals.
10. Take charge and focus attention on yourself.
11. Create and innovate – open new possibilities, provoke thought in others.
12. Catalyse productive and accountable thinking, conversation, and action.

The reason behind asking questions depends on the situation when the question has been asked. Let us enumerate some obvious situations:

- Ask the audience a verbal question, e.g. a teacher questioning pupils.
- Ask a person a verbal question but in front of an audience, e.g. at a scientific conference.
- Ask a person a written question but in front of an audience, e.g. at a scientific conference.
- Ask a person an anonymous written question but in front of an audience, e.g. at a company's meeting.
- Ask a "face to face" question.
- Ask a "face to face" question on the phone.
- Ask an audience a written question, e.g. through the internet.
- Ask a personal written question, e.g. through email.
- Ask an SMS question. It is completely different in comparison to the previous point because of size, convenience and money.
- Ask the Artificial Intelligent Question-Answering System (AIQAS) a verbal question.
- Ask the AIQAS a MQ. This is exactly our case and just two reasons (1 and 6) among the 12 listed earlier are appropriate for this particular situation.

On the one hand two reasons to send an MQ have been selected, on the other hand there should be some external, implicit reason that causes the user to send an MQ. Let us call such a reason a meta-reason.

Meta-reason for Mobile Query Sending

It is very important to underline that in this paper we consider the precise situation when an MQ is sent to the AIQAS but not to another person. Initially, we decided to find out about the SMS-activities during the year i.e. it to discover whether and how SMS-activities depend on seasons, weekdays, holidays, events etc. (see Figure 1). Red (top) charts represent quantity of daily SMS. The set of events is represented as events hierarchical structure, nodes of which, in essence, are clusters and have links to corresponding nodes of clusters structure. Suppose an MQ was sent on Saturday, 08.07.06. MQAS will extract a list of events taking place on or around this day from the KB (see Figure 1) and create a list of potentially active clusters: for example, *football, tennis, motor racing, music,* and *show*. Thus, MQAS has selected five WDAD ahead of the MQ analysis. Of course, for

MQ: *"What tablets can help me with headache?"* these WDAD would not be used but they might help to disambiguate MQ related to one of these enumerated clusters.

Figure 1. SMS-activities and Events

Mobile Queries Classification

As part of its MQ processing component, the MQAS first attempt to classify each MQ by their type and represented as a pair: <Question Type> ↦ <Answer Type> [15]. Only two kinds of question are allowed for MQ: (1) Specific question (*what, who, why, how, where, when*, etc.) and (2) Yes-No question. For example, instead of asking: *"Name all directors of 2ergo"*, the specific question *Who* should be used: *"Who are directors of 2ergo?"*. Examples of MQ classification are shown below:

- What ↦ (money, definition, name, etc): "What was the monetary value of the Nobel Peace Price in 1989?"
- What ↦ (person, organisation): "What costume designer decided that M. Jackson should wear only one glove?"
- What ↦ (date): "In what year did Ireland elect its first woman president?"
- What ↦ (location): "What is the capital of the Ukraine?"
- Which ↦ (person, organisation): "Which former Klu Klux Klan member won an elected office in the U.S.?"
- Which ↦ (date): "In which year was New Zealand excluded from the ANZUS alliance?"
- Which ↦ (location): "Which city has the oldest relationship as sister-city with Los Angeles?"
- Who ↦ (person, organisation): "Who is the author of the book "The Iron Lady: A Biography of Margaret Thatcher"?"
- Whom ↦ (person, organisation): "Whom did Italy beat in the final of 2006 Football World Cup?"
- Where ↦ (location): "Where is Kharkov?"
- When ↦ (date): "When did the Jurassic Period end?"
- Why ↦ (reason): "Why do people shake hands to show friendliness?"
- How ↦ (action): "How did Socrates die?"
- How many ↦ (number): "How many James Bond novels are there?"
- How much ↦ (money, price, time, etc.): "How much pizza do Americans eat in a day?"
- How long ↦ (time, distance): "How long does it take to travel from Plymouth to Rawtenstall?"

As for a Yes-No MQ, in general, it is better to avoid asking a Yes-No MQ unless one wants to receive a yes or no answers. For example, instead of asking *“Are there any tablets to relieve my headache?”* it is better to ask *“What tablets can relieve my headache?”*. A Yes-No MQ highlights a very important problem of communication with both a “human” QAS or with AIQAS. By default, when a user sends an MQ he/she does not doubt that the recipient is a human being. That is why the Yes-No MQ like *“Does anyone know ... ?”, “Can someone ... ?“, “Is it possible to ... ?”*, etc is very popular. In a call centre the recipients appear to find such MQ annoying [16] — and are likely to return logically impeccable but dismissive answers like *“Yes, you can be helped”* in response to the MQ *“Can you help me to find an address of 2ergo?”*. As for our MQAS, on the one hand, we decided to give the impression that our system is not artificial, on the other hand, MQAS is always “well behaved” (i.e. it is simply impossible to annoy it – since it is a machine) and that is why MQAS would reply to user *“Yes, you can be helped but it is better to ask a specific question”*.

The MQ classifier is critical during the stage of an MQ’s type recognition (see Figure 2). A list of supplementary key words for each *question word* allows the specification of the type of expected response more precisely (see an example of *When* description in Knowledge Base Structure session).

Derivation of implicit information for MQ Response

The primary focus of MQAS is for the creation of a coherent, understandable answer that is responsive to the originally posed MQ. The factually explicit MQ e.g. *“What is the Taj Mahal?”* does not require a great deal of effort to create the answer if the MQAS KB contains the fact: *“Taj Mahal is a beautiful mausoleum at Agra built by the Mogul emperor Shah Jahan (completed in 1649) in memory of his favourite wife”*. But the problems are, firstly, it is impossible beforehand to classify MQ as explicit or implicit, and secondly, in reality most of the MQs are implicit. For example, if the MQAS KB contains the fact: *“Aleksandr Sergeyevich Pushkin is a Russian poet (1799 -1837)”*. The obvious implicit MQs related to this fact are *“When was Pushkin born?”* or *“When did Pushkin die?”*. Below we discuss the possible ways for the derivation of implicit information from the KB.

In this paper we continue to apply a ***psycholinguistic approach*** to natural language (NL) processing [1]. The only system truly capable of adequate understanding of an MQ is human. What is more, children seem to use NL effortlessly in spite of *not knowing the grammar*. Parsing was taught in school as an algorithmic task. For derivation of implicit information a ***natural inference engine*** based on *human reasoning* is used. *Human reasoning* might be described using fuzzy attributes: *approximate, common sense, default, enumerative, evidential, hypothetical, inexact, integrating, plausible, procedural, taxonomic.* In other words, instead of logically making some conclusion a human would very often ***justify*** the decision based on maximum argumentation about problem solving within his/her knowledge. At present it is understandable that the real knowledge base (KB) is ***incomplete***, ***inconsistent, inaccurate*** and ***open***. These facts do not necessarily permit the use of a traditional logical approach. For example, assume that the following two statements are in the KB: (1) *“A student likes to read a detective story”,* (2) *“A student does not like to read mathematical books”* and the MQ simply states *“Does a student like to read?”*.

The derivation of implicit information is not always based on the logical operations:

- From the conditional judgement *"If 'A' then 'B'"* it is not possible to draw the conclusion *"If not 'A' then not 'B'"*. As for the sentence, *"If you smoke you will fall ill"* the conclusion *"If you do not smoke you will not fall ill"* is not true since non-smokers also become ill. Representation of the conditional judgements stated in the linguistic form and the conclusions drawn from them depend widely on their contents.
- The examinees were given the sentences: *"Acorns always grow on an oak"* and *"Acorns grow on a tree"*, they concluded on this basis: *"This tree is an oak"* [14]. The potential fallacy of this conclusion becomes evident when considering another example apparently having the same structure. From the sentences, *"All football-players are good runners"* and *"Peter is a good runner,"* one can hardly draw the conclusion that *"Peter is a football-player"*. In the first instance, at first glance, the rule of inadmissibility of the general use judgement was violated. It is not possible to conclude that *"All B's are A's"* from the judgement *"All A's are B's"*. It should be specified that the simple transformation of the general judgement is not always

incorrect but only in those cases when the participants of the judgement have dissimilar size. In the case of the identical sizes the simple transformation of a general judgement is possible. For instance, the judgement *"All squares are equilateral rectangles"* transforms into the judgement *"All equilateral rectangles are squares"* as the "participants": *"square"* and *"equilateral rectangle"* are of equal size. Really, there are no rectangles other than equilateral ones being squares. As to the judgement *"All football-players are good runners"* this is not the case. Here the "participants" sizes: *"football-players"* and *"good runners"* are different and the simple transformation of this judgement is inadmissible.

Cognitive transformations as the natural inference source for the derivation of implicit information:

- Taking into account the cognitive transformations based on the semantic relations featuring the properties of symmetry, transitivity and additivity allows MQAS to answer a Yes-No type MQ. For example:
  (1) Fact: *"'A' is a colleague of 'B'". MQ: "Is 'B' a colleague of 'A'?".*
  (2) Fact: *"'A' buys a doll".* MQ: *"Does 'A' buy a toy?".*
  (3) Fact: *"'A' is a teacher of 'B'".* MQ: *"Is 'B' a pupil of 'A'?".*
- It is important to process facts on the things belonging to transformation. For example, from the fact: *"Mary has given a book to John after New Year"* three different MQ might be answered:
  (1) *"Did John have this book before New Year?"*;
  (2) *"Does John has this book now?"*;
  (3) *"Does Mary have it now?".*

The derivation of implicit information based on *human reasoning* is a part of Reply Search Engine (RSE). The central question to be addressed by any QAS is how the storage of information is organised in KB and we now turn to consider this.

Knowledge Base Structure

In the general case, under the *knowledge base structure* one should understand the regularity of data distribution in memory assuring the storage of various links between separate elements of stored information. At every moment KB deals only with relatively *small fragments* of the external world. So, the corresponding structures are needed to integrate these fragments separated in time into the integral picture. The structures obtained as a result of integration should contain more information than had been used for its creation. The organisation of knowledge storage should make allowance for such features of the human memory as [17]:

- associativity;
- ability to reflect similar features for different objects and different features for similar objects;
- hierarchical and heterarchical organisation of information. The idea of heterarchical approach correlates naturally with the human ability to use all kinds of information in the process of natural language understanding. As Quillain remarks "*... a full association structure... forms simply large, very complicated net of nodes and unidirectional memory links between them... The predetermined hierarchy of "super-" and "subclasses" is absent; every word is a "patriarch" in its own hierarchy if some process of search initiates with it. Analogously every word is in different places within hierarchies of large diversity of wordy concepts if the search process starts with them*" [18, p.5];
- associative relations weight variable;
- representation of the environment statistical properties;
- independence of the knowledge extraction time from the volume of knowledge being stored in this memory;
- the knowledge cannot realistically be regarded as a static resource, to be accumulated and stored within a system. It is a formative, *self-organising* character, with the ability to change the organisation within which it is held.

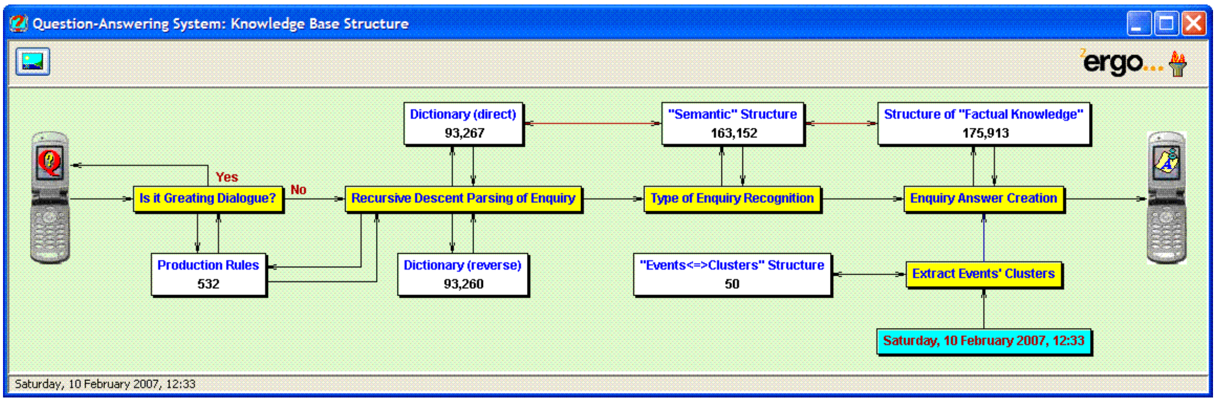

Figure 2. Knowledge Base Structure

The attempt to take into account these features of human memory in MQAS KB has been undertaken. The structure of KB is shown in Figure 2. MQAS's consists of six different structures:

- Direct and Reverse Dictionary (DD and RD respectively) are represented by an L-Tree structure. DD is an initial structure which provides the recognition of *new* words, the normalisation of *known* words and the direct links with the corresponding nodes of Semantic structure and Factual Knowledge structure. RD together with DD is used to correct wrong words, spelling them automatically.
- Semantics structure (or Sequential-Simultaneous Structure (SSS)) provides the *sequential* and *simultaneous* analysis of string information and handling of *new* and/or *known* sentences or combinations of words. SSS is a special combination of hierarchical and network structures. Each of the elements in SSS are associated with several other elements logically including itself or included by it. With the successive presentation of facts as a sequence of words the strongest relation in it is the relation between the nearest neighbouring words. *"Their succeeding one after another presents evidently an important condition of structuring"* [19, p.231]. The ***sequential*** part of SSS provides ***hierarchical*** organisation of information and the ***simultaneous - heterarchical*** organisation of information
- Events-Clusters Structure (ECS)
- Factual Knowledge Structure (FKS)
- Production Rules Structure (PRS) [4]. New class of PR Question has been added. The example of PR in format: <Class> ↦ <Antecedent> ⇒ <Consequent> is shown below.

Question ↦ When ⇒ 0:born,start,begin,commence,come,become;1:live,interval,period;
2:die,end,finish,stop;3:occur,happen,find.

Let us come back to a fact: *"Aleksandr Sergeyevich Pushkin is a Russian poet (1799 -1837)"* in FKS. Assume the MQ *"When was Pushkin born?"* has been asked. In the consequent of PR *Question ↦ When* subclass 0 has been found, which means that the left-hand number from the interval should be selected as a birth year. First of all MQAS tries to find in FKS the direct fact for an answer like *"Pushkin was born in 1799"*. If such a fact does not exist MQAS starts to analyse facts with date's interval and then from the appropriate fact select left number as a birth year i.e. *"In 1799"*.It is worth noting that, MQAS always try to minimise the length of the response because of the display constraints of the mobile device and the SMS message.

At the end of this session it is important to emphasise:

- Local KB of any MQAS cannot be complete, in other words, we shall never be able to establish information completeness of KB.
- MQAS should be self-learning i.e. when MQAS cannot find a suitable response in the local KB it should search using the Internet and then not only send the reply to the questioner but also automatically extend the local KB. The searching of information across the Internet represents the greatest problem because the result may identify a huge set of documents among which the appropriate response will need to be found, We shall discuss this problem in the next paper.

Mobile Query Parsing

The main requirement for MQAS is to reply to any (even non-standard or poorly formed but, nevertheless, meaningful) user's MQ. But it is not easy to find the answer even for an ideal MQ because for an artificial intelligent system like MQAS the power of natural language to describe the same events but in quite different ways is a great problem. For example, the primitive action: *"take by theft"* might be described as: *"hook", "snitch", "thieve", "cop", "knock off", or "glom".* The main purpose of MQ processing is to understand ***what was meant*** rather than ***what was said***. The mechanism of query parsing is very simple: *"eliminating the unnecessary until only the necessary remain"* and has been discussed elsewhere [2]. Here we just remind ourselves of the main steps involved in MQ processing. This MQAS takes the MQ as a character sequence, locates the MQ boundaries, and converts the original MQ to a *skeleton*. Such conversion will require several steps:

- MQ related synonyms' dictionary creation to equate synonymous words and phrases, such as "NLP" and "Natural Language Processing";
- Irregular verb normalisation. Once the word has been identified then it should be changed back to its simplest form for efficient word recognition. For example, *writes, writing, wrote, written* will be changed to *write* and the corresponding attributes of the original form will be saved;
- Noisy (non-searchable) word elimination;
- Plural to singular conversion.

The skeleton of the MQ is matched against all relevant data in SSS to find the appropriate links to FKS. The result of such searching is shown on Figure 3.

Figure 3. The examples of MQAS responses

Conclusion

The possible solutions of how to provide the response to a question if the application domain is unknown have been considered. In the result of that, the MQAS effectively places information directly into the hands of any users - eliminating the need for technical support specialists continually to address *ad hoc* requests from end users.

The MQAS are addressed from both scientific and industrial perspectives. Whether MQAS is searching the local KB, or the worldwide web, MQAS understands the relationships between words, enabling it to extract all key concepts and automatically build a semantic index organised in a problem-solution format. Because MQAS extracts and organises the content, the user receives specific and relevant answers to his/her MQ — not a list of documents.

Bibliography

[1] G.Coles, T.Coles, V.A.Lovitskii, "Natural Interface Language", Proc. of the VIII-th International Conference on Knowledge-Dialogue-Solution: KDS-99, Kacivelli (Ukraine), 104 -109, 1999.

[2] T.Coles, V.A.Lovitskii, "Text Searching and Mining", J. of Artificial Intelligence, National Academy of Sciences of Ukraine, Vol 3, 488-496, 2000.

[3] D.Burns, R.Fallon, P.Lewis, V.Lovitskii, S.Owen, "Verbal Dialogue Versus Written Dialogue", International Journal "Information Theories and Application", Vol 12(4), 369-377, 2005.

[4] G.Francis, M.Lishman, V.Lovitskii, D.Traynor, "Instantaneous Database Access", Proc. of the XII-th International Conference on Knowledge-Dialogue-Solution: KDS-2006, Varna (Bulgaria), 314 - 322, 2006.

[5] K.Braithwaite, M.Lishman, V.Lovitskii, D.Traynor, "Distinctive Features of Mobile Messages Processing", Proc. of the XII-th International Conference on Knowledge-Dialogue-Solution: KDS-2006, Varna (Bulgaria),322 - 329, 2006.

[6] www.hsbc.com.sg/1/2/personal/services/sms-from-hsbc

[7] www.2ergo.com (2safeguard™, 2NC™)

[8] www.revenue.ie

[9] www.smseducation.org

[10] www.edinburghairport.com

[11] Google SMS query service

[12] www.Think6Results.com

[13] M.Adams, Inquiry Institute, www.inquiryinstitute.com

[14] R.J.Harris, G.E.Monaco, "Psychology of Pragmatic Implication: Information Processing between the Lines", J. Exp. Psychol. General, 107, 1-22, 1978.

[15] D.Moldovan, S.Harabagiu, et al. "Lasso: A Tool for Surfing the Answer Net.", TREC-8 Draft Proceedings, NIST, 65-73, 1999.

[16] www.82ask.com.

[17] V.A.Lovitskii and K.Wittamore, "AMONIL: Achievement of Multiple Objectives using a Natural Interface Language", Proc. of the International Joint Conference on Knowledge-Dialogue-Solution: KDS-97, Yalta (Ukraine), 270-281, 1997.

[18] M.R.Quillian, "Word concepts: A theory and simulation of some basic semantic capabilities", C.I.P. working paper 79, Carnegie Inst. of Technology, Pittsburgh, 1965.

[19] J.Hoffmann, "Das Aktive Gedachtnis. Psycologische Experimente und Theorien zur Menschlichen Gedachtnistatigkeit", VEB Deutscher Vergal der Wissenschaften, Berlin, 1982.

Authors' Information

Vladimir Lovitskii *– 2 Ergo Ltd, St. Mary's Chambers, Haslingden Road, Rawtenstall, Lancashire, BB4 6QX, UK, vladimir@2ergo.com*

Michael Thrasher *– University of Plymouth, Plymouth, Devon, PL4 6DX, UK, mthrasher@plymouth.ac.uk*

David Traynor *– 2 Ergo Ltd, St. Mary's Chambers, Haslingden Road, Rawtenstall, Lancashire, BB4 6QX, UK, david.traynor@2ergo.com*