Title: METHOD AND APPARATUS FOR BUILDING GRAMMARS WITH LEXICAL SEMANTIC CLUSTERING IN A SPEECH RECOGNIZER

Abstract: A method and system for building a grammar module for a speech application. The method includes the step of clustering phrases having a semantic similarity. The grammar module comprises phrases in a machine-readable format and semantic concepts associated with the phrases. According to another aspect, the grammar module includes embedded semantic interpretations associated with the semantic concepts.
For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
The present invention relates to speech recognition systems, and most particularly to a method and system for building grammars with lexical semantic clustering.

Automated speech applications allow a person to interact with a computer-implemented system using their voice and ears in much the same manner as interacting with another person. Such systems utilize automated speech recognition technology, which interprets the spoken words from a person and translates them into a form, which is semantically meaningful to a computer, for example, strings or other types of digital data or information.

With current known speech recognition technology or speech recognizers, the domain of discourse needs to be sufficiently small to achieve practical recognition rates. Speech applications are typically modeled as a sequence of questions and responses, i.e. the system poses a question, and the person (i.e. user) provides a response. Furthermore, the questions or prompts are typically worded so as to restrict the domain of discourse and elicit from the user a response that the system is capable of recognizing. This model is required because a speech recognizer only understands words or phrases that it has been programmed \textit{apriori} to understand.
It is known in the art to program a speech recognizer with a context free grammar. The context free grammar comprises a precise specification of the recognized phraseology. In a typical speech application, the context free grammar represents the designer's best prediction of what a person will say in response to a particular question or prompt posed by the system. When the scope of reasonable or expected responses to a question is sufficiently small, a context free grammar can be provided which successfully predicts the spoken responses that will be made by all the system's users. However, as the phraseology expands, for example, with an open-ended question, it becomes increasingly difficult to predict a priori all the responses and variations that will be provided by a user.

On one level, the speech recognizer attempts to emulate the human ability to understand language. However, the speech recognizer has no ability to understand natural language as the human brain can. The speech recognizer simply executes computer code that identifies phonemes in the digitized sound wave generated by a person's voice and then attempts to find a corresponding phrase in the provided grammar that has a similar sequence of phonemes. It is typically the responsibility of the speech application to associate a semantic meaning to the results of the speech recognizer. And in many cases, the associated semantics are manually determined.

The design of a context free grammar for a speech application typically involves two design considerations. The first design consideration comprises predicting phraseology that encompasses all the possible responses that may be given by a user to the questions or prompts posed by the speech application. The second design consideration comprises providing a semantic interpretation or mapping for each possible response, i.e. word or phrase, that may be provided by a user of the system. It will be appreciated that the design of a system with open-ended questions presents particular challenges because the large number of responses makes it very
difficult to program \textit{a priori} all or even most of the phraseology for the possible responses. It also becomes very difficult to determine \textit{a priori} the set of semantic interpretations for mapping the phraseology or phrases corresponding to the responses. Furthermore, when semantics interpretations are manually associated with phrases, the shear number of phrases makes this task time consuming, error prone, and costly.

[0007] Accordingly, it will be appreciated that there remains a need for improvements in the art.

SUMMARY OF THE INVENTION

[0008] The present invention provides a method and system for creating a grammar module suitable for a speech application.

[0009] According to one aspect, the grammar module includes one or more semantic concepts. The semantic concepts are generated by clustering semantically similar phrases into groups, wherein each of the clustered phrases represents the same or a similar semantic concept.

[0010] In a first embodiment, the present invention provides a method for creating a grammar module for use with a speech application, the method comprises the steps of: collecting phrases associated with one or more voice responses; transcribing the collected phrases into a machine-readable format; clustering selected ones of the collected phrases into one or more semantic concepts, and wherein the selected collected phrases corresponding to each of the semantic concepts have a related meaning; building a grammar module based on the collected phrases and the semantic concepts.
In another embodiment, the present invention provides a system for building a grammar module for a speech application, the system comprises: means for collecting phrases associated with one or more voice responses; means for transcribing the collected phrases into a machine-readable format; means for clustering selected ones of the collected phrases into one or more corresponding semantic concepts; and means for creating a grammar module based on the collected phrases and the semantic concepts.

In a further embodiment, the present invention provides a method for generating a grammar module for a speech application, the method comprises the steps of: collecting one or more phrases associated with one or more voice responses; transcribing the collected phrases into a machine-readable format; clustering selected ones of the collected phrases into one or more semantic concepts, and wherein the selected collected phrases in each of the semantic concepts have a similar meaning; interpreting at least some of the semantic concepts; building a grammar module based on the collected phrases, the semantic concepts and the interpreted semantic concepts.

Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of embodiments of the invention in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Reference will now be made to the accompanying drawings which show, by way of example, embodiments of the present invention, and in which:
Fig. 1 shows in diagrammatic form a networked communication system incorporating a voice recognition mechanism according to an embodiment of the present invention;

Fig. 2 shows in flowchart form a method for building a grammar module according to an embodiment of the present invention; and

Fig. 3 shows in flowchart form a method for building a grammar module according to another embodiment of the present invention.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Reference is first made to Fig. 1, which shows in diagrammatic form a voice based communication system 100 incorporating a speech recognition mechanism and techniques according to the present invention. As shown, the voice based communication system 100 comprises a telecommunication network 110 and a voice application 120. The telecommunication network 110 may comprise, for example, a public or a private telephone or voice network or a combination thereof. The voice application 120 in the context of the following description comprises a voice node 130 and a speech application server 140. The speech application server 140 runs or executes a speech application 142, e.g. a standalone computer program or software module or code component or function. The voice node 130 includes a speech recognizer indicated generally by reference 132. The speech recognizer 132 comprises a software module or engine which converts voice signals or speech samples into digital data or other forms of data which are recognized by the speech application server 140, and in the other direction, the speech recognizer 132 converts the digital data or voice information generated by the speech application 142 into vocalizations or other types of audible signals. As indicated in Fig. 1, the speech
recognizer 132 includes a grammar module according to an embodiment of the invention and indicated generally by reference 150.

[00019] In the context of a speech application, a large number or sample of spoken answers are typically empirically collected for each question that is or may be posed by the application. The phrases are collected from a population that is representative of the users of the speech application. In the accompanying description, the collection of phrases, typically tens of thousands in number, is called or termed phraseology. In a speech application, the phraseology is typically dominated by phrases that are in-context; i.e. phrases that comprise on-topic responses for the question posed by the application. However, most speech applications are designed to accommodate a statistically significant number of phrases that are out-of-context. Out-of-context phrases are not consistent with the question posed, but in the larger context of the speech application, may still have some relevance. As will be described in more detail below, embodiments of the present invention provide a mechanism or process for building a grammar module for the speech application which can accommodate both in-context and out-of-context phrases and which includes lexical clustering according to an aspect of the invention.

[00020] Users or subscribers use telecommunication devices, for example, a fixed line telephone set 112, or wireless or cellular communication devices 114, to communicate with each other via the telecommunication network 110 by dialing the directory number or DN associated with another user’s telephone. The voice node 130 is also assigned a directory number and a user dials the directory number of the voice node 120 to initiate a call session with the speech application running 142 on the speech application server 140. The speech application 142 may comprise, for example, a business listings directory accessed by voice commands. The voice node 130 handles the call from the telephone set 112 or the communication device 114 of a user, and the speech recognizer 132 handles the conversion of voice signals, e.g.
commands and responses to voice prompts, into digital data or other form of
information which is recognizable to the speech application 142. The speech
application server 140, in turn, controls or handles the call session. During the call
session, the speech application 142 running on the server 140 will typically execute
several dialog forms. For example, the speech application 142 prompts the user with
one or more questions, waits for a response from the user, and then provides further
prompts or processing, as dictated by the particular application. The speech
recognizer 132 converts the prompts generated by the speech application 142 into
corresponding vocalizations or other types of voice or audible signals. The speech
recognizer 132 converts the responses provided by the user into corresponding digital
data. As will be described in more detail below, the grammar module 150 is utilized
by the speech recognizer 132 and provides a mechanism for building a grammar base
or module for use by the speech application 142.

[00021] The speech recognizer 132 and speech application 142 are implemented as
software on the voice node 130 and the speech application server 140, respectively,
and may comprise a standalone computer program, a component of software in a
larger program, or a plurality of separate programs, or software, hardware, firmware
or any combination thereof. The particular details or programming specifics for
implementing software, computer programs or computer code components or
functions for performing the operations or functions associated with the embodiments
of the present invention will be readily understood by those skilled in the art. While
described in the context of a voice-based networked communication system, it will be
appreciated that the present invention has wider applicability and is suitable for other
types of voice-based or speech recognition applications.

[00022] Reference is next made to Fig. 2, which shows in flowchart form a
method 200 according to one embodiment of the invention for creating or generating
a grammar module, for example, the grammar module 150 (Fig. 1) for the speech
application 142 running on the speech application server 140 (Fig. 1). As described
above with reference to Fig. 1, a user of the speech application 142 initiates a call
from a telecommunications device, for example, a cellular phone 114, over the
telecommunication (e.g. a public or private telephone) network 110. The voice node
130 and the speech recognizer 132 handle the call from the user, and the speech
application server 140 handles the call session. During the call session, the speech
application 142 executes several dialog forms, which include prompting the user, i.e.
calling party, with a question, and then listening for the caller's response. The
responses or replies received from the caller are handled by the speech recognizer
132, which utilizes the grammar module 150. As will be described in more detail
below, the process according to an embodiment of the invention provides for the
creation of the grammar module 150 comprising semantic concepts and context free
grammars for open-ended questions, i.e. questions that can have a large number of
distinct responses. For example, in a speech accessible business directory, the
question "what type of business are you looking for" can result in 10,000 or more
distinct responses.

[00023] As shown in Fig. 2, the first step indicated by block 210 involves the
collection of a large number or sample of spoken responses. The spoken responses are
typically collected from a population that is statistically representative of the
population that will be using the speech application 142 (Fig. 1). In general, the
environment in which the phrases are collected will accurately simulate the
anticipated environment of the speech application. In addition, the words and sentence
structure chosen by a person to respond to a question can depend on several
environmental factors, including, but not limited to: the time of day; the
communication medium; the person's location; and, perhaps most significantly, the
knowledge that the person's conversational partner is an automated computer system.
The next step in the process 200, indicated by reference 220, comprises transcribing the collected phrases to text or some other digitized form. The collected and transcribed phrases are saved in a digital transcription file 222, which is stored as part of a database or in computer memory, for example, in the voice node 130 (Fig. 1) or the speech application server 140 (Fig. 1). The next step indicated by block 230 comprises clustering the phrases from the transcription file 222. According to this aspect, a computer-implemented clustering process or algorithm is applied to the transcription phrases in the file 222 to cluster semantically similar phrases into groups called semantic concepts. For example, the phrases *my car needs gasoline* and *my auto requires petrol* belong to the same semantic concept, because they have the same semantic meaning. In the context of the present description, the clustering algorithm or process provides *lexical semantic clustering*, and according to one embodiment, the clustering algorithm may be implemented as described by the following pseudo code:

\[
\begin{align*}
C & \leftarrow \{ \} \\
& \text{for each phrase } p \text{ do} \\
& \quad \text{if } C = \{ \} \text{ then} \\
& \quad \quad C \leftarrow C \cup \{ \{p\} \} \\
& \quad \text{else} \\
& \quad \quad c' \leftarrow \text{argmax}_5(p, c) \\
& \quad \quad \quad \text{if } S(p, c') > t \text{ then} \\
& \quad \quad \quad \quad c' \leftarrow c' \cup \{p\} \\
& \quad \quad \text{else} \\
& \quad \quad \quad C \leftarrow C \cup \{i, p\} \\
\end{align*}
\]
With reference to the pseudo code, the lexical semantic clustering algorithm starts or begins by initializing the set of semantic concepts \( C \) to an empty set. Next, each phrase is compared to the semantic concepts in \( C \). Because \( C \) is initially empty, the first phrase always begins a new semantic concept, which is added to the semantic concepts set \( C \). For each subsequent phrase \( p \), the phrase \( p \) is compared to each semantic concept to the find the semantic concept whose phrases are most similar to the phrase \( p \). The function \( S \) computes the similarity between a phrase and a semantic concept, as described in more detail below. If the similarity between the phrase \( p \) is greater than a threshold \( \kappa \), then the phrase \( p \) is added to the semantic concept; otherwise, the phrase \( p \) becomes the seed of a new semantic concept. The algorithm terminates or ends when all of the transcribed phrases have been analyzed, at which point \( C \) contains the set of semantic concepts. The set of semantic concepts \( C \) are stored in a digital semantic concepts file 232, e.g. a phrase clusters file. In other words, the semantic concepts \( C \) comprise a set of semantically equivalent phrases. In accordance with this aspect, the meaning or relevance of the semantic concept is typically determined by the context of the application.

[00025] An aspect of the clustering operation in step 230 as described above involves quantitatively measuring the similarity between two phrases. Known methods for measuring similarity typically incorporate some form of vectorization of the phrases. The vocabulary size of the phraseology determines the dimension of the vector or vector space. For example, a phraseology comprising \( N \) distinct words results in an \( N \) dimensional space with each word being represented by a dimension. Furthermore, a particular phrase is represented by a vector having non-zero components for each word in the phrase. For example, the phrase \textit{coffee shop} is represented as \((0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)\), where the two 1's correspond to the words \textit{coffee} and \textit{shop}, and the 0's correspond to the words in the phraseology, but not in the phrase \textit{coffee shop}. In typical speech applications, the vocabulary size is often several thousand, so the phrase vector is dominated by 0 components.
In the above example for a vector-based implementation, each component has either the value 0 or 1, indicating either the absence or the presence of a word in a phrase. It will be appreciated that this scheme has the disadvantage of treating all words with equal importance. According to another aspect, the concept of information content can be applied to the vectorization of each phrase, wherein the 0's remain, and for each word in a phrase, the corresponding vector component is assigned the information content of the word. The information content for a word $w$ is $-\log_2 P(yw)$, where $P(w)$ is the probability of the word $w$ occurring. The simplest estimate for $P(w)$ is $\frac{H_w}{N}$, where $H_w$ is the number of phrases containing the word $w$ and $N$ is the number of phrases. According to another aspect, more complex probability models, for example, using n-grams and Bayes' Theorem, may be applied.

The similarity between two vectorized phrases $x$ and $y$ can be determined using the Jaccard similarity coefficient:

$$s(x, y) = \frac{x \cdot y}{\|x\|^2 + \|y\|^2 - x \cdot y}$$

or the cosine measure:

$$s(x, y) = \frac{x \cdot y}{\|x\| \|y\|}$$

It will also be appreciated that notwithstanding a finding of dissimilarity, for example, using Jaccard's coefficient or the cosine similarity measurements described above, phrases can still be semantically similar. For example, the phrases *my car needs gasoline* and *my auto requires petrol* are semantically similar, but because these two exemplary phrases have few words in
common, the similarity measurements, i.e. Jaccard's or cosine, fail to identify the similarity. To address this potential occurrence during vectorization of phrases, the clustering operation provides for the interjection of synonymous terms. For example, for the phrase *my car needs gasoline*, the terms *auto* and *petrol* are inserted into the phrase vector, as synonyms for the words *car* and *gasoline*. The injected synonyms will typically have the same vector weight as the original word or term. According to another aspect, hypernyms and/or hyponyms are inserted into the phrase vector. The injected terms will have a scaled weight which is less than the original term, because the injected terms have related, but not equivalent, semantics.

[00029] The vectorization process can be improved further by applying a *word sense* tag or indicator for each word according to another embodiment. For example, the word *glasses* can mean a container used for drinking, or eyewear. The word sense tag indicates which meaning of a word is intended. Depending on the context in which the word is being used, the word sense tag may be determined manually or algorithmically (e.g. through the execution of a computer program, function or code component). There may also be instances where a word sense tag cannot be determined, for example, where there is ambiguity in the entire phrase. According to this aspect, each word, or most words, in the phrase are tagged with a word sense.

When vectorizing a phrase, words with different senses are considered distinct, and if a word is determined to be ambiguous, then in the vector form, each word sense is represented by a non-zero component.

[00030] Reference is made back to Fig. 2, and the clustering algorithm in step 230. The clustering operation includes determining the similarity between the phrases and the semantic concepts by performing a similarity measurement, for example, a scalar similarity measure. According to one embodiment, the similarity between the phrase *p* and the semantic concept *c* (i.e. represented as a set of phrases), is defined as follows:
The clustering operation 230, and execution of the clustering algorithm, yields a set of semantic concepts, which are stored in a semantic concepts file indicated by reference 232. Next in step 240, the process 200 uses the semantic concepts file 232 to build a grammar file or module 242 for the speech recognizer (i.e. the speech recognizer 132 in Fig. 1). The grammar module 242 (i.e. indicated by reference 150 in Fig. 1) comprises a machine-readable format and is used by the speech recognizer 132 to recognize or decode words and phrases in the responses provided by the user (for example, as described above), and the decoded speech is then provided to the speech application 142 (Fig. 1) for further processing according to the application.

[00031] Reference is next made to Fig. 3, which shows in flowchart form a process 300 according to another embodiment for creating or generating a grammar module for a speech application, for example, as described above for Fig. 1. The process 300 is similar to the process 200 of Fig. 2, and includes a collect phrases operation (step 310), a transcribe phrases operation (step 320), creation of a transcription file (reference 322), a cluster phrases operation (step 330) and creation of a semantics file (reference 332). The process 300 performs or executes these operations in a manner similar to that described above for the process 200 of Fig. 2.

[00032] As shown in Fig. 3, the process 300 according to this embodiment includes a semantic interpretation operation in step 340. The semantic interpretation step 340 operates to create a semantic interpretation for each semantic concept C, and the semantic interpretations are stored in a file denoted by reference 342. The semantic interpretation operation in step 340 typically comprises a manual process,
which is performed by a person skilled in the appropriate domain. The build grammar
operation in step 350 builds a machine-readable grammar file 352. The grammar file
352 also includes the semantic interpretations which are converted to a machine-
readable format and embedded with the grammar elements. The implementations
details associated with this operation will be within the understanding of one skilled
in the art.

[00033] In summary, the processes and clustering algorithm according to the
present invention allows semantically equivalent phrases to be grouped together,
which in turn provides the capability to organize and identify distinct semantic
concepts present in the phraseology of interest or relevant to a particular speech
application. When the phraseology is sufficiently large, and the semantic
interpretations are determined using a manual process, the creation of semantic
concepts can greatly reduce the manual effort because semantic interpretations need
only to be done for each semantic concept, and not every phrase.

[00034] The present invention may be embodied in other specific forms
without departing from the spirit or essential characteristics thereof. Certain
adaptations and modifications of the invention will be obvious to those skilled in the
art. Therefore, the presently discussed embodiments are considered to be illustrative
and not restrictive, the scope of the invention being indicated by the appended claims
rather than the foregoing description, and all changes which come within the meaning
and range of equivalency of the claims are therefore intended to be embraced therein.
WHAT IS CLAIMED IS:

1. A method for creating a grammar module for a speech application, said method comprising the steps of:

   - collecting phrases associated with one or more voice responses;
   - transcribing said collected phrases into a machine-readable format;
   - clustering selected ones of said collected phrases into one or more semantic concepts, and wherein said selected collected phrases in each of said semantic concepts have a related meaning;

   building a grammar module based on said collected phrases and said semantic concepts.

2. The method as claimed in claim 1, wherein said step of clustering comprises the step of identifying one or more words in each of said collected phrases and associated said collected phrases with a semantic concept when one or more of said words have a meaning which is similar or the same.

3. The method as claimed in claim 2, wherein said step of identifying one or more words comprises generating a vector for said collected phrase, said vector having an element for each of a plurality of words in said collected phrase, and comparing the vector for said collected phrase to a vector for one of said semantic concepts, and associating said collected phrase with said semantic concept if said vector has a number of elements exceeding a predefined threshold.

4. The method as claimed in claim 3, wherein said step of building a grammar module comprises converting a plurality of grammar elements into a machine-
readable format and converting said semantic concepts into a machine-readable format, and storing said machine-readable grammar elements and semantic concepts in a computer file.

5. The method as claimed in claim 3, wherein one or more of said vector elements includes an indicator, said indicator providing information about said associated vector element.

6. The method as claimed in claim 5, wherein said indicator comprises a content indicator providing a probability indicator for the occurrence of a word.

7. The method as claimed in claim 5, wherein said indicator comprises a word sense indicator providing an intended meaning for a word.

8. The method as claimed in claim 3, further including the step of inserting one or more synonymous terms for one or more words in said collected phrases wherein said one or more words have a synonymous term, and said vector including a corresponding element for at least some of said synonymous terms.

9. The method as claimed in claim 3, further including the step of inserting one or more hypernyms into said vector, and said one or more hypernyms having a weighting.

10. A system for building a grammar module for a speech application, said system comprising:
means for collecting phrases associated with one or more of said voice
responses;

means for transcribing said collected phrases into a machine-readable format;

means for clustering selected ones of said collected phrases into a plurality of
semantic concepts, wherein each of said semantic concepts comprises one or more
collected phrases having a similar meaning;

means for creating a grammar module based on said collected phrases and said
semantic concepts.

11. The system as claimed in claim 10, wherein said means for clustering includes
means for characterizing each of said selected collected phrases as a vector, said
vector having one or more elements corresponding to one or more words comprising
said collected phrase, and each of said semantic concepts including one or more
vectors having an element for each of a plurality of words associated with said
semantic concept.

12. The system as claimed in claim 11, further including means for comparing
each of said collected phrase vectors to one or more of said semantic concept vectors
based on a similarity measure, and means for grouping one or more of said collected
phrases when said similarity measure exceeds a predetermined threshold.

13. The system as claimed in claim 12, further including means for inserting one
or more synonymous terms for one or more words in said collected phrases wherein
said one or more words have a synonymous term, and said vector including a
corresponding element for at least some of said synonymous terms.
14. The system as claimed in claim 12, further including means for inserting one or more hypernyms into said vector, and said one or more hypernyms each having an associated weighting.

15. A method for creating a grammar module suitable for use with a speech application, said method comprising the steps of:

   collecting phrases associated with one or more voice responses;

   transcribing said collected phrases into a machine-readable format;

   grouping one or more of said collected phrases into a plurality of groups, wherein each of said groups has an associated semantic, said one or more collected phrases being grouped based on a similarity between said collected phrase and the associated semantic concept for said group; and

   building a grammar module based on said collected phrases and said semantic concepts.

16. The method as claimed in claim 15, wherein said step of grouping comprises determining a similarity between said collected phrase and the associated semantic concept for said group, and comparing said similarity to a predefined threshold, and adding said collected phrase to the group associated with said semantic concept if said predefined threshold is satisfied.

17. The method as claimed in claim 16, further including the step utilizing said collected phrase not satisfying said predefined threshold for a new semantic concept.

18. The method as claimed in claim 17, wherein said semantic concepts comprise a plurality of semantically equivalent words or phrases.
19. The method as claimed in claim 16, wherein said similarity is determined according to a similarity function.

20. A method for generating a grammar module for a speech application, said method comprising the steps of:

   collecting one or more phrases associated with one or more voice responses;

   transcribing said collected phrases into a machine-readable format;

   clustering selected ones of said collected phrases into one or more semantic concepts, and wherein said selected collected phrases in each of said semantic concepts have a similar meaning;

   interpreting at least some of said semantic concepts;

   building a grammar module based on said collected phrases, said semantic concepts and said interpreted semantic concepts.

21. The method as claimed in claim 20, wherein said step of building a grammar module comprises creating a machine-readable grammar file.

22. The method as claimed in claim 21, further including converting said interpreted semantic concepts into a machine-readable format and embedding said interpreted semantic concepts in said machine-readable grammar file.

23. The method as claimed in claim 20, wherein said step of interpreting each of said semantic concepts comprises converting said interpreted semantic concepts into a machine-readable format.
INTERNATIONAL SEARCH REPORT

A.  CLASSIFICATION OF SUBJECT MATTER
IPC: GIOL 15/00 (2006.01), GIOL 15/06 (2006.01), GIOL 15/08 (2006.01)
According to International Patent Classification (IPC) or to both national classification and PCT

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC s: GIOL 15/*

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
Delphion database, Derwent database, WEST database, IEEE database, Google Patent database, Canadian Patent Database. Keywords: grammar/grammar module/context free grammar, speech/semantic*, lexical semantic clustering/cluster/group/phrases,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 6,317,707 B1 (Bangalore et al.) 13 November 2001 (13-1 1-2001)</td>
<td>1, 2, 10, 11, 15, 20-23</td>
</tr>
<tr>
<td></td>
<td><em>column 1, line 6 to column 4, line 22</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>claim 1</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>abstract</em></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>US 5,860,063 (Gorin et al.) 12 January 1999 (12-01-1999)</td>
<td>1, 2, 10, 11, 15, 20-23</td>
</tr>
<tr>
<td></td>
<td><em>column 1, line 5 to column 2, line 29</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>column 3, lines 34-43</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>column 4, lines 26-30</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>column 6, lines 55-66</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>column 8, line 58 to column 9, line 19</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>claim 1</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>abstract</em></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[X] Further documents are listed in the continuation of Box C.  [X] See patent family annex.

* Special categories of cited documents
‘A’ document defining the general state of the art which is not considered to be of particular relevance
‘E’ earlier application or patent but published on or after the international filing date
‘L’ document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
‘O’ document referring to an oral disclosure, use, exhibition or other means
‘P’ document published prior to the international filing date but later than the priority date claimed

‘T’ later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
‘X’ document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
‘Y’ document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

Date of the actual completion of the international search
10 May 2007 (10-05-2007)

Date of mailing of the international search report
17 August 2007 (17-08-2007)

Name and mailing address of the ISA/CA
Canadian Intellectual Property Office
Place du Portage 1, Cl 14 - 1st Floor, Box PCT 50 Victoria Street
Gatineau, Quebec K1A 0C9
Facsimile No : 001-819-953-2476

Authorized officer
C. Fletcher (819-934-7564)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 6,415,248 B1 (Bangalore et al) 2 July 2002 (02-07-2002)</td>
<td>1-23</td>
</tr>
<tr>
<td>A</td>
<td>US 5,794,193 (Gorm) 11 August 1998 (11-08-1998)</td>
<td>1-23</td>
</tr>
<tr>
<td>A</td>
<td>WO 00/73936 A1 (Ehsam et al) 7 December 2000 (07-12-2000)</td>
<td>1-23</td>
</tr>
<tr>
<td>Patent Document Cited in Search Report</td>
<td>Publication Date</td>
<td>Patent Family Member(s)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2003033338 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2004064303 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE69818161 D1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE69818161 T2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP0890942 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES2210623 T3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP1 1143491 A</td>
</tr>
<tr>
<td>US6173261</td>
<td>09-01-2001</td>
<td>NONE</td>
</tr>
<tr>
<td>US64155248</td>
<td>02-07-2002</td>
<td>NONE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP0763817 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP9167161 A</td>
</tr>
<tr>
<td>WO0073936</td>
<td>07-12-2000</td>
<td>AUS451800 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP1224569 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP2003505578 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2002032564 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2002128921 A1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US2004199375 A1</td>
</tr>
</tbody>
</table>