

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 March 2008 (27.03.2008)

PCT

(10) International Publication Number
WO 2008/036879 A2

(51) International Patent Classification:
G10L 15/18 (2006.01) *G10L 17/00* (2006.01)

(74) Agents: SUSTEIN, Bruce, D. et al.; Bromberg & Sunstein LLP, 125 Summer Street, Boston, MA 02110-1618 (US).

(21) International Application Number:
PCT/US2007/079131

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) International Filing Date:
21 September 2007 (21.09.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/826,400 21 September 2006 (21.09.2006) US

(71) Applicant (*for all designated States except US*): NUANCE COMMUNICATIONS, INC. [US/US]; 1 Wayside Road, Suite 100, Burlington, MA 01803-4613 (US).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,

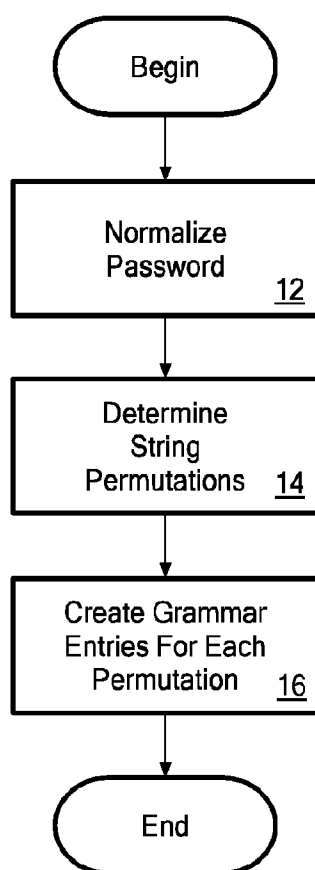
(72) Inventor; and

(75) Inventor/Applicant (*for US only*): BREUER, Richard [DE/DE]; Leo-Meurer-Weg 5, 52477 Alsdorf (DE).

[Continued on next page]

(54) Title: GRAMMAR GENERATION FOR PASSWORD RECOGNITION

(57) Abstract: A password grammar for speech recognition is described. A password is normalized into a list of strings of a plurality of character types such as letters and numerals. For each string of letters, one or more corresponding letter permutations are determined which represent pronounceable combinations of that string. Then, for each letter permutation, a corresponding recognition grammar entry is created for a speech recognition grammar.



WO 2008/036879 A2



PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished
upon receipt of that report*

Attorney Docket: 2639/B62WO

Grammar Generation for Password Recognition

[0001] This application claims priority from U.S. Provisional Patent Application 60/826,400, filed September 21, 2006, the contents of which are incorporated herein by reference.

[0002] Field of the Invention

[0003] The invention generally relates to automatic speech recognition, and more specifically, to a speech recognition grammar for spoken passwords.

Background Art

[0004] Many computer applications require the user to provide a difficult to predict password, such as Hitt-222X. Typically, such a password must be typed in via a keyboard input device. Although speech recognition is becoming more commonly used for text input, speech recognition of such difficult to predict passwords is more difficult than for normal words.

Summary of the Invention

[0005] Embodiments of the present invention are directed to creating a password grammar for speech recognition. A password is normalized into a list of strings of different character types including letters and numerals. For each string of letters, one or more corresponding letter permutations are determined that represent pronounceable combinations of that string. Then for each letter permutation, a corresponding recognition grammar entry is created for a speech recognition grammar.

[0006] In further embodiments, for each string of numerals, one or more corresponding number permutations are also determined which represent pronounceable combinations of that string. And for each number permutation, a corresponding recognition grammar entry is also created for the speech recognition grammar. Similarly, in some embodiments, for each string of non-alpha-numeric characters, one or more corresponding character

permutations are also determined which represent pronounceable combinations of that string. And for each character permutation, a corresponding recognition grammar entry is also created for the speech recognition grammar.

[0007] The grammar may be, for example, a Speech Recognition Grammar Specification (SRGS) grammar.

Brief Description of the Drawings

[0008] Figure 1 shows a flow chart of major functional steps for creating a speech recognition password grammar according to an embodiment of the present invention.

[0009] Figure 2 shows a normalized listing of a password according to an embodiment.

[0010] Figure 3 shows a permutation listing of a password according to an embodiment.

Detailed Description of Specific Embodiments

[0011] Embodiments of the present invention are directed to speech recognition techniques for analyzing a user password and creating a grammar for speech recognition of the password. Rejection of wrong passwords is based on the recognizer's features to reject out-of-grammar utterances. In the following discussion, we consider the sample password "Hitt-222X" and a closed Speech Recognition Grammar Specification (SRGS) grammar is created. Of course, this is simply an illustrative example which should not be understood to limit the scope of the invention.

[0012] Figure 1 shows a flow chart of major functional steps for creating a speech recognition password grammar according to one specific embodiment of the present invention. Initially, an existing password is normalized into a list of strings containing numeral strings or digits, strings of letters or single letters, and special characters, step 12. For example, as shown in Fig. 2, the password **20** Hitt-222X is normalized to a letter string **21** with a value of Hitt **22**, a special string -, a numeral string 222, and another letter string X.

[0013] Next, all the characters sequences—letters, numbers, and non-alphanumeric characters—are processed by creating permutations of all possible and meaningful substrings so as to determine string permutations, step 14. This is to create grammar paths containing all pronounceable substrings. More specifically, all possible permutations 30 are created as shown in Fig. 3. So for example, the letter string 31 Hitt 32 is permuted into all possible variations of that string, Hit t, Hi tt, Hi t t, etc. Then all letter permutations are removed that contain sequences longer than one letter without a vowel, which do not consist of all the same character, assuming these cannot / will not be spoken. In the example, nothing will change, since the combinations “H i tt” and “Hi t t” which would be removed, contain “tt”, a sequence of the letter “t”. Therefore they are not removed.

[0014] Once the string permutations have been developed, a grammar entry is created for each remaining permutation, step 16. For example, in one specific embodiment, the grammar contains one rule “password”. This rule contains a <one-of> for all alternative paths through the password. Each <one-of> itself is a sequence of tokens. Each token is created from one entry in the result list. For each non-alphanumeric character, a reference to a matching predefined rule is created, e.g. <ruleref uri = ”#hyphen”/>. If the first letter is lowercase, a <ruleref uri=”#lower”/> is created, otherwise <ruleref uri=”#upper”/>.

#lower and #upper are predefined rules, e.g.:

```
<rule id="upper">
    <item repeat="0-1">upper case</item>
</rule>

<rule id="lower">
    <item repeat="0-1">lower case</item>
</rule>
```

If the sequence does not consist of all the same letter, the (lowercase) sequence of letters itself is put in the rule. Otherwise, if the letters occur two times or more, a “double x” sequence is created, etc. In applicable cases, a “<number> times x” is created. If the number is lower than a maximum (999.999), a natural number is created. Otherwise this alternative is omitted, knowing that shorter sequences are available as well. Additionally, sequences of the same digit are treated specially. If one digit occurs two times, a “double x” sequence is created. In any case “<number> times x” is created.

[0015] Embodiments of the invention may be implemented in any conventional computer programming language. For example, preferred embodiments may be implemented in a procedural programming language (*e.g.*, C) or an object oriented programming language (*e.g.*, C++ or Perl). Alternative embodiments of the invention may be implemented as pre-programmed hardware elements, other related components, or as a combination of hardware and software components.

[0016] Embodiments can be implemented as a computer program product for use with a computer system. Such implementation may include a series of computer instructions fixed either on a tangible medium, such as a computer readable medium (*e.g.*, a diskette, CD-ROM, ROM, or fixed disk) or transmittable to a computer system, via a modem or other interface device, such as a communications adapter connected to a network over a medium. The medium may be either a tangible medium (*e.g.*, optical or analog communications lines) or a medium implemented with wireless techniques (*e.g.*, microwave, infrared or other transmission techniques). The series of computer instructions embodies all or part of the functionality previously described herein with respect to the system. Those skilled in the art should appreciate that such computer instructions can be written in a number of programming languages for use with many computer architectures or operating systems. Furthermore, such instructions may be stored in any memory device, such as semiconductor, magnetic, optical or other memory devices, and may be transmitted using any communications technology, such as optical, infrared, microwave, or other transmission technologies. It is expected that such a computer program product may be distributed as a removable medium with accompanying printed or electronic documentation (*e.g.*, shrink wrapped software), preloaded with a computer system (*e.g.*, on system ROM or fixed disk), or distributed from a server or electronic bulletin board over the network (*e.g.*, the Internet or World Wide Web). Of course, some embodiments of the invention may be implemented as a combination of both software (*e.g.*, a computer program product) and hardware. Still other embodiments of the invention are implemented as entirely hardware, or entirely software (*e.g.*, a computer program product).

[0017] Although various exemplary embodiments of the invention have been disclosed, it should be apparent to those skilled in the art that various changes and modifications can be made which will achieve some of the advantages of the invention without departing from the true scope of the invention.

What is claimed is:

1. A method for creating a password grammar for speech recognition comprising:
normalizing a password into a list of strings of a plurality of character types, the
character types including letters and numerals;
for each string of letters, determining one or more corresponding letter permutations
representing pronounceable combinations of the string; and
for each letter permutation, creating a corresponding recognition grammar entry for a
speech recognition grammar.
2. A method according to claim 1, further comprising:
for each string of numerals, determining one or more corresponding number
permutations representing pronounceable combinations of the string; and
for each number permutation, creating a corresponding recognition grammar entry for
the speech recognition grammar.
3. A method according to claim 1, further comprising:
for each string of non-alphanumeric characters, determining one or more
corresponding character permutations representing pronounceable
combinations of the string; and
for each character permutation, creating a corresponding recognition grammar entry
for the speech recognition grammar.
4. A method according to claim 1, wherein the grammar is a Speech Recognition
Grammar Specification (SRGS) grammar.
5. A speech recognition grammar for password recognition comprising:
a recognition rule for each pronounceable permutation of each substring of letters.
6. A speech recognition grammar according to claim 5, further comprising:
a recognition rule for each pronounceable permutation of each substring of numerals.

7. A speech recognition grammar according to claim 5, further comprising:
a recognition rule for each pronounceable permutation of each substring of non-alphanumeric characters.
8. A speech recognition grammar according to claim 5, wherein the grammar is a Speech Recognition Grammar Specification (SRGS) grammar.
9. A computer program product for creating a password grammar for speech recognition, the product comprising:
program code for normalizing a password into a list of strings of a plurality of character types, the character types including letters and numerals;
program code for determining, for each string of letters, one or more corresponding letter permutations representing pronounceable combinations of the string; and
program code for creating, for each letter permutation, a corresponding recognition grammar entry for a speech recognition grammar.
10. A computer program product according to claim 9, further comprising:
program code for determining, for each string of numerals, one or more corresponding number permutations representing pronounceable combinations of the string;
and
program code for creating, for each number permutation, a corresponding recognition grammar entry for the speech recognition grammar.
11. A computer program product according to claim 1, further comprising:
program code for determining, for each string of non-alphanumeric characters, one or more corresponding character permutations representing pronounceable combinations of the string; and
program code for creating, for each character permutation, a corresponding recognition grammar entry for the speech recognition grammar.
12. A computer program product according to claim 1, wherein the grammar is a Speech Recognition Grammar Specification (SRGS) grammar.

1/1

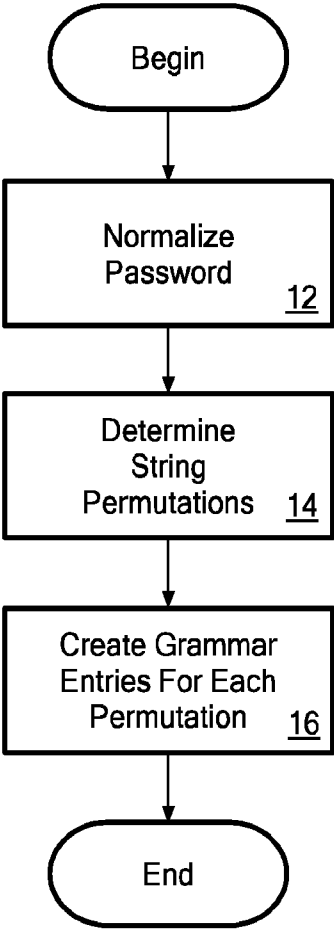


FIG. 1

20

Flag	Contents
Letter	Hitt
Special	-
Numeral	222
Letter	X

21 22

FIG. 2

30

Flag	Contents
Letter	Hitt
	Hit t
	Hi tt
	Hi t t
	H itt
	H it t
	H i tt
	H i t t
Numeral	222
	22 2
	2 22
	2 2 2
Letter	X

31 32

FIG. 3