17712 U.S. PTO
083105

PTO/SB/16 (12-04)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

113277 U.S.PTO
60/712412
083105

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. _____

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
ROBERT A.	KENNEWICK	Seattle, WA

Additional inventors are being named on the _____ *separately numbered sheets attached hereto*

TITLE OF THE INVENTION (500 characters max)

DYNAMIC SPEECH SHARPENING

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

[X] The address corresponding to Customer Number: 00909

OR

[ ] Firm or Individual Name

Address

Address

City	State	Zip
Country	Telephone	Fax

ENCLOSED APPLICATION PARTS *(check all that apply)*

[X] Application Data Sheet. See 37 CFR 1.76

[X] Specification *Number of Pages* 13

[X] Drawing(s) *Number of Sheets* 3

[ ] CD(s), Number _____

[ ] Other (specify) _____

Application Size Fee: If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

TOTAL FEE AMOUNT ($): 100.00

[X] Applicant claims small entity status. See 37 CFR 1.27.

[ ] A check or money order is enclosed to cover the filing fee and application size fee (if applicable).

[ ] Payment by credit card. Form PTO-2038 is attached.

[X] The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: 033975. A duplicative copy of this form is enclosed for fee processing.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

[X] No.

[ ] Yes, the name of the U.S. Government agency and the Government contract number are: _____

SIGNATURE _____

Date August 31, 2005

TYPED or PRINTED NAME Sean L. Ingram

REGISTRATION NO. 48,283 *(If appropriate)*

Docket Number: 090611-0312471

TELEPHONE (703) 905.2107

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

MICROSOFT CORP.
EXHIBIT 1004

DYNAMIC SPEECH SHARPENING

FIELD OF THE INVENTION

The invention is related to enhancing automated speech interpretation.

BACKGROUND OF THE INVENTION

The field of automated speech interpretation is in increasingly higher demand. One use of automated speech interpretation is as an input for electronic devices. This may enable a user to simply speak to an electronic device rather than inputting commands, or other information, manually, by uploading, or by other input methods. Controlling various electronic devices through speech may enable the user to use the electronic devices more efficiently.

However, existing technology in the field of automated speech interpretation, such as standard speech engines and other devices, are not able to interpret natural human speech with a suitable accuracy to sufficiently control some electronic devices. Shortcomings in accuracy may have many causes, including poor microphones, extraneous noises, unclear speech by the user, an accent of the user, a quality of standard search engines, and other causes.

In light of these and other problems, there is a need for enhanced automated speech interpretation that may interpret natural human speech with an augmented accuracy.

SUMMARY

One aspect of the invention may relate to a system for enhancing automated speech interpretation. The system may be used to sharpen a speech interpretation made by a speech engine. The system may be modeled at least partially after one or more speech pattern recognition techniques used by humans. A speech pattern recognition technique may include interpreting speech using words, a word sequence, a word combination, a word position, a context, a phonetic similarity between two or more words, a part of speech, or other techniques.

In some embodiments of the invention, the system may receive information related to a verbalization made by a user. The system may receive this information from a speech engine. The information may include textual information, or other information, that may represent a plurality of preliminary interpretations of the verbalization made by the user. The plurality of preliminary interpretations may be arranged in a predetermined form, such as an array, a matrix, or other forms.

According to various aspects of the invention, the system may include a policy module. The policy module may manage and/or provide one or more policies that may enable the system to designate one of the preliminary interpretations as a probable interpretation of the verbalization made by the user. A policy may include one or more policy parameters for distinguishing between the preliminary interpretations to designate a probable interpretation. A policy parameter may include a policy vocabulary, a word position in the verbalization, a word combination, a sentence structure, or other parameters. A policy may include a parameter weighting scheme. The parameter weighting scheme may weight individual policy parameters according to one or more weighting factors, such as, a frequency of use, a difficulty to understand, or other factors.

The policy module may include a profile policy sub-module. The profile policy sub-module may manage and/or provide one or more profile policies. A profile policy may correspond to a user. The profile policy may include one or more profile policy parameters that may be tailored to the user. The profile policy may be used as a base policy in interpreting any verbalizations made by the user. The profile policy may be augmented to enable the system to provide more accurate interpretations of verbalizations made by the user. The augmentation may include a user provided augmentation, such as, providing additional vocabulary (e.g. the names in an address book, etc.), providing a personalized pronunciation, or other user provided augmentations. The augmentation may include a non-user provided augmentation, such as, updates generated by a third party (e.g. a commercial administration and/or maintenance entity), or other non-user provided augmentations. The augmentation may include an automated

augmentation, such as alterations made to a profile parameter weighting scheme in an automated manner due to past speech patterns of the user, or other automated augmentations. Other augmentations may be made.

The policy module may include a context policy sub-module. The context policy sub-module may manage and/or provide one or more context policies. A context policy may include one or more context policy parameters that may be tailored to a verbalization context. The context policy may enhance an ability of the system related to interpreting verbalizations made by the user in the verbalization context corresponding to the context policy. A context policy may be augmented to enable the system to provide more accurate interpretations of verbalizations made in a verbalization context corresponding the context policy. The augmentation may include a user provided augmentation, a non-user provided augmentation, an automated augmentation, or other augmentations.

According to various embodiments of the invention, the policy module may include a context tracking sub-module. The context tracking sub-module may track the verbalization context of a consecutive series of verbalizations by the user. The context tracking sub-module may utilize one or more conversation trees to track the series of verbalizations. The context tracking sub-module may track one or more past verbalization contexts of the series of verbalizations, may track one or more current verbalization contexts of the series of verbalizations, and/or may make predictions regarding one or more future verbalization contexts of the series of verbalizations. The policy module may utilize information about the verbalization context of the series of verbalizations generated by the context tracking sub-module to manage and/or provide one or more context policies.

In some embodiments of the invention, the system may include an interpretation sharpening module. The interpretation sharpening module may generate a probable interpretation of a verbalization made by a user. The interpretation sharpening module may generate the probable interpretation based at least in part on a plurality of preliminary interpretations of the verbalization and one or more policies. The plurality of preliminary interpretations may be provided by the speech engine. The policies may be provided by the policy module.

The interpretation sharpening module may apply one or more policies, such as, one or more context policies and/or a profile policy, to the preliminary interpretations to generate one or more interpretation scores. An interpretation score may be related to the likelihood of a particular preliminary interpretation being a correct interpretation of the verbalization. The preliminary interpretation corresponding to the highest (or lowest) interpretation score may then be designated as a probable interpretation of the verbalization.

According to various embodiments of the invention, the system may include an interpretation history analysis module. The interpretation history analysis module may enable the system to augment one or more policies based on an analysis of past interpretations related to previously interpreted vocalizations. The augmentations enabled by the interpretation history analysis module may include an automated augmentation, or other augmentations.

The interpretation history analysis module may include an information storage sub-module. The information storage sub-module may store interpretation information related to past verbalizations, such as, one or more preliminary interpretations associated with a past verbalization, one or more interpretation scores associated with a past verbalization, one or more probable interpretations associated with a past verbalization, whether or not a past verbalization was interpreted correctly, or other information.

The interpretation history analysis module may include a frequency sub-module. The frequency sub-module may use some or all of the information stored in the information storage sub-module to generate one or more frequencies related to one or more past verbalizations. For example, the frequency sub-module may calculate a word usage frequency, a word combination frequency, a frequency related to a set of verbalizations that are phonetically similar but have distinct meanings, an interpretation error frequency for a particular verbalization, or other frequencies.

Information stored and/or generated by the interpretation history analysis module may be used to augment the system. In some embodiments, one or more frequencies generated may be used to augment one or more policies. For

example, one or more policy parameters and/or a policy weighting scheme of a policy may be augmented based on a frequency generated by the interpretation history analysis module. Other augmentations using information stored and/or generated by the interpretation history analysis module may be made.

In some embodiments of the invention, the system may include a policy agent handler. The policy agent handler may augment, update, remove, and/or provide one or more policy agents to the system. A policy agent may include one or more context policies. The policy agent handler may be controlled, directly or indirectly, by a third party (e.g. a commercial entity, etc.). The policy agent handler may augment, update, remove, and/or provide the policy agents to the system as part of a commercial agreement, such as a licensing agreement, a subscription agreement, a maintenance agreement, or other agreements.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates an exemplary embodiment of a system for enhancing automated speech interpretation.

FIG. 2 illustrates an exemplary embodiment of a system for enhancing automated speech interpretation.

FIG. 3 illustrates a flow chart of an exemplary embodiment of a method for enhancing automated speech interpretation.

DETAILED DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates an exemplary embodiment of a system 110 for enhancing automated speech interpretation. System 110 may receive information related to a verbalization made by a user. System 110 may receive this information from a speech engine 112. The information may include textual information, or other information, that may represent a plurality of preliminary interpretations of the verbalization. The plurality of preliminary interpretations may be arranged in a predetermined form, such as an array, a matrix, or other forms.

According to various aspects of the invention, system 110 may include a policy module 114. Policy module 114 may manage and/or provide one or more policies that may enable system 110 to designate one of the preliminary interpretations as a probable interpretation of the verbalization. A policy may include one or more policy parameters for distinguishing between the preliminary interpretations to designate a probable interpretation. A policy parameter may include a policy vocabulary, a word position in the verbalization, a word combination, a sentence structure, or other parameters. A policy may include a parameter weighting scheme. The parameter weighting scheme may weight individual policy parameters according to one or more weighting factors, such as, a frequency of use, a difficulty to understand, or other factors.

In some embodiments of the invention, system 110 may include an interpretation sharpening module 116. Interpretation sharpening module 116 may generate a probable interpretation of the verbalization. Interpretation sharpening module 116 may generate the probable interpretation based at least in part on the plurality of preliminary interpretations. The policies may be provided by policy module 114.

According to various embodiments of the invention, system 110 may include an interpretation history analysis module 118. Interpretation history analysis module 118 may enable system 110 to augment one or more policies based on an analysis of past interpretations related to previously interpreted vocalizations. The augmentations enabled by interpretation history analysis module 118 may include an automated augmentation, or other augmentations.

In some embodiments of the invention, system 110 may include a policy agent handler 120. Policy agent handler 120 may augment, update, remove, and/or provide one or more policy agents to system 110. A policy agent may include one or more context policies

FIG. 2 illustrates an exemplary embodiment of system 110 including policy module 114. Policy module 114 may include a profile policy sub-module 210. Profile policy sub-module 210 may manage and/or provide one or more profile policies. A profile policy may correspond to a user. The profile policy may

include one or more profile policy parameters that may be tailored to the user. The profile policy may be used as a base policy in interpreting any verbalizations made by the user. Profile policy sub-module 210 may enable an augmentation of the profile policy. Augmentation of the profile policy may enable system 110 to provide more accurate interpretations of verbalizations made by the user. An augmentation may include a user provided augmentation, such as, providing additional vocabulary (e.g. the names in an address book, etc.), providing a personalized pronunciation, or other user provided augmentations. An augmentation may include a non-user provided augmentation, such as, updates generated by a third party (e.g. a commercial administration and/or maintenance entity), or other non-user provided augmentations. An augmentation may include an automated augmentation, such as alterations made to a profile parameter weighting scheme in an automated manner due to past speech patterns of the user, or other automated augmentations. Other augmentations may be made.

Policy module 114 may include a context policy sub-module 212. Context policy sub-module 212 may manage and/or provide one or more context policies. A context policy may include one or more context policy parameters that may be tailored to a verbalization context. The context policy may enhance an ability of the system related to interpreting verbalizations made by the user in the verbalization context corresponding to the context policy. Context policy sub-module 212 may enable an augmentation of a context policy. Augmentation of a context policy may enable system 110 to provide more accurate interpretations of verbalizations made in a verbalization context corresponding the context policy. An augmentation may include a user provided augmentation, a non-user provided augmentation, an automated augmentation, or other augmentations.

According to various embodiments of the invention, policy module 114 may include a context tracking sub-module 214. Context tracking sub-module 214 may track the verbalization context of a consecutive series of verbalizations by the user. Context tracking sub-module 214 may utilize one or more conversation trees to track the series of verbalizations. Context tracking sub-module 214 may track one or more past verbalization contexts of the series of

verbalizations, may track one or more current verbalization contexts of the series of verbalizations, and/or may make predictions regarding one or more future verbalization contexts of the series of verbalizations. Policy module 114 may utilize information about the verbalization context, generated by context tracking sub-module 214, to manage and/or provide one or more context policies.

In some embodiments, interpretation sharpening module 116 may apply one or more policies to the preliminary interpretations provided by speech engine 112. The policies may include one or more context policies, a profile policy, or other policies. The policies may be provided by policy module 114. Applying the policies to the preliminary interpretations may generate one or more interpretation scores. An interpretation score may be related to the likelihood of a particular preliminary interpretation being a correct interpretation of the verbalization. The preliminary interpretation corresponding to the highest (or lowest) interpretation score may then be designated by interpretation sharpening module 116 as a probable interpretation of the verbalization.

According to various embodiments of the invention, interpretation history analysis module 118 may include an information storage sub-module 216. Information storage sub-module 216 may store interpretation information related to past verbalizations, such as, one or more preliminary interpretations associated with a past verbalization, one or more interpretation scores associated with a past verbalization, one or more probable interpretations associated with a past verbalization, whether or not a past verbalization was interpreted correctly, or other information.

Interpretation history analysis module 118 may include a frequency sub-module 218. Frequency sub-module 218 may use some or all of the information stored in information storage sub-module 216 to generate one or more frequencies related to one or more past verbalizations. For example, frequency sub-module 218 may calculate a word usage frequency, a word combination frequency, a frequency related to a set of verbalizations that are phonetically similar but have distinct meanings, an interpretation error frequency for a particular verbalization, or other frequencies.

Information stored and/or generated by interpretation history analysis module 118 may be used to augment system 110. In some embodiments, one or more frequencies generated, may be used to augment one or more policies. For example, one or more policy parameters and/or a policy weighting scheme of a policy may be augmented based on a frequency generated by interpretation history analysis module 118. Other augmentations related to information stored on and/or generated by interpretation history analysis module 118 may be made.

Policy agent handler 120 may augment, update, remove, and/or provide one or more policy agents to system 110. A policy agent may include one or more context policies. Policy agent handler 120 may be controlled, directly or indirectly, by a third party (e.g. a commercial entity, etc.). Policy agent handler 120 may augment, update, remove, and/or provide the policy agents to system 110 as part of a commercial agreement, such as a licensing agreement, a subscription agreement, a maintenance agreement, or other agreements.

FIG. 3 illustrates a flow chart of an exemplary embodiment of a method 310 for enhancing automated speech interpretation. Method 310 may include a verbalization operation 312. At verbalization operation 312 a verbalization of a user may be electronically captured. The verbalization may be captured by a microphone or other electronic audio capture device.

According to various embodiments of the invention, method 310 may include a preliminary interpretation operation 314. At preliminary interpretation operation 314 one or more preliminary interpretations of the verbalization may be made. The preliminary interpretations may be represented in a textual format, or otherwise represented. The preliminary interpretations may be arranged in a predetermined form, such as, an array, a matrix, or other forms. Preliminary interpretation operation 314 may be performed by speech engine 112.

The preliminary interpretations may be provided to system 110 at a provide preliminary interpretations operation 316. The preliminary interpretations may be received in system 110 by interpretation sharpening module 116.

In some embodiments of the invention, method 310 may include a provide policies operation 318. Provide policies operation 318 may be performed

concurrently with provide preliminary interpretations operation 316. At provide policies operation 318 one or more policies may be selected and/or provided to system 110. The policies selected and/or provided may include a profile policy, a context policy, or other policies. The policies may be selected and/or provided by policy module 114. The policies may be provided to interpretation sharpening module 116.

At an apply policies operation 320, one or more policies may be applied to the preliminary interpretations. The policies applied to the preliminary interpretations may include the policies provided at provide policies operation 318, or other policies. Applying the policies to the preliminary interpretations may generate one or more interpretation scores. An interpretation score may be related to a likelihood of a particular preliminary interpretation being a correct interpretation of the verbalization. The policies may be applied to the preliminary interpretations by interpretation sharpening module 116.

According to various embodiments of the invention, at a determine probable interpretation operation 322 a highest (or lowest) interpretation score may be determined, and a preliminary interpretation corresponding to the highest (or lowest) interpretation score may be designated as a probable interpretation. The probable interpretation may represent an enhanced interpretation of the verbalization of the user. Determine probable interpretation operation 322 may be performed by interpretation sharpening module 116.

Method 310 may include a store interpretation information operation 324. At store interpretation information operation 324, system 110 may store interpretation information related to the verbalization. Interpretation information may include one or more preliminary interpretations associated with a past verbalization, one or more interpretation scores associated with a past verbalization, one or more probable interpretations associated with the verbalization, whether or not the verbalization was interpreted correctly, or other information. The interpretation information may be stored at interpretation history analysis module 118.

In some embodiments of the invention, some or all of the interpretation information stored at store interpretation information operation 324 may be used to determine one or more frequencies at a determine frequencies operation 326. The frequencies calculated at determine frequencies operation 326 may include one or more frequencies related to past verbalizations, such as, a word usage frequency, a word combination frequency, a frequency related to a set of verbalizations that are phonetically similar but have distinct meanings, an interpretation error frequency for a particular verbalization, or other frequencies. Determine frequencies operation 326 may be performed by interpretation history analysis module 118.

In various embodiment, system 110 may decide whether an augmentation may be made at an augmentation decision operation 328. The decision concerning system augmentation may be based at least in part on information generated at determine frequencies block 326, such as one or more frequencies, or other information. If system 110 decides that augmentation may not be made, no further action is taken until another verbalization is captured.

In some instances, system 110 may determine that augmentation may be made and control of system 110 may transfer to an augment system operation 330. Augment system operation 330 may include making an augmentation to system 110. For example, one or more policies may be augmented in accordance with information generated by determine frequencies operation 326. The policies augmented may include a context policy, a profile policy, or other policies.

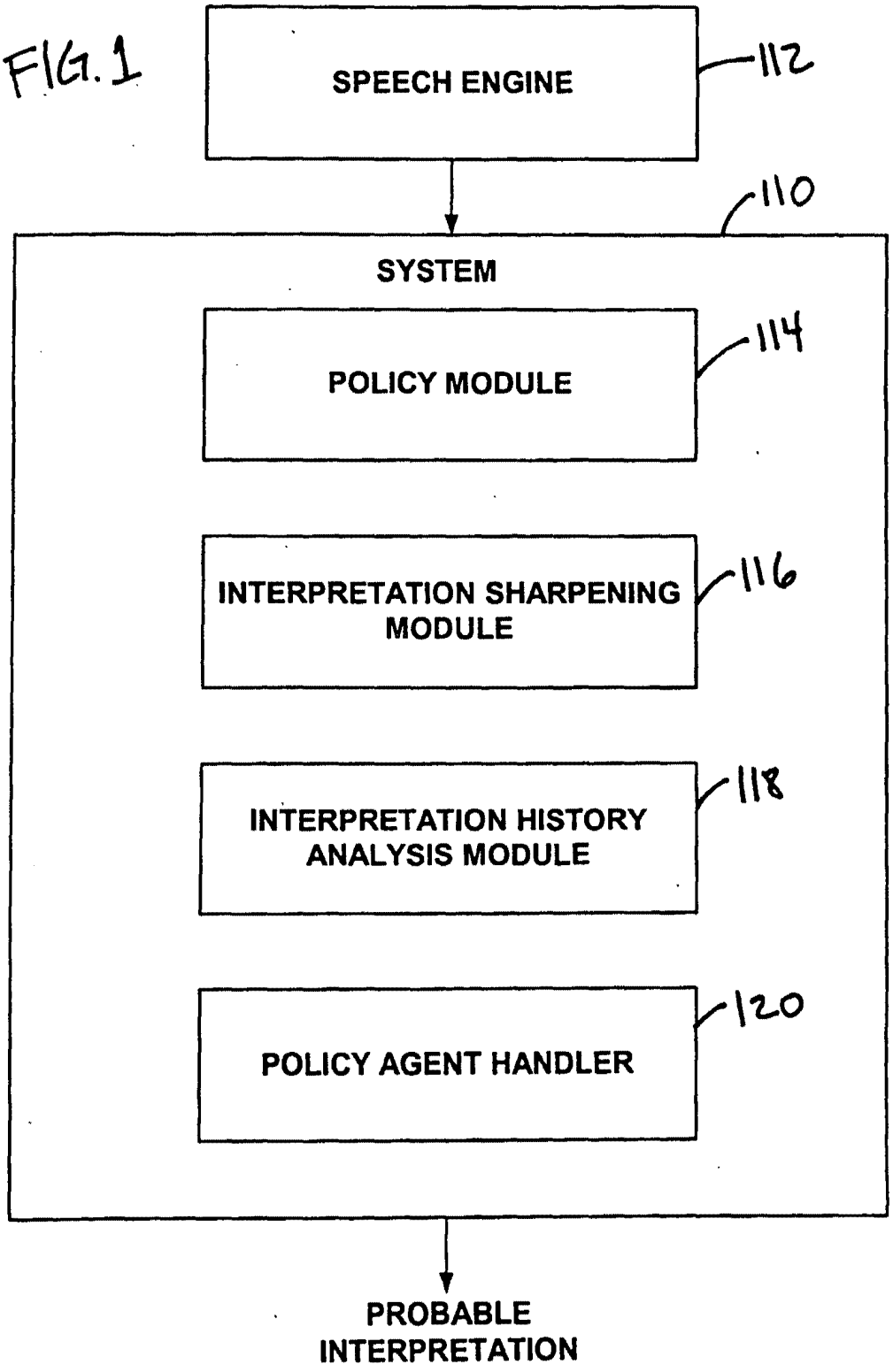
What is claimed is:

THIS PAGE BLANK (USPTO)

ABSTRACT

A system and method for enhancing automated speech interpretation. The system and method may be used to sharpen a speech interpretation made by a speech engine. The system and method may be modeled at least partially after speech pattern recognition techniques used by humans. A speech pattern recognition technique may include interpreting speech using words, a word sequence, a word combination, a word position, a context, a phonetic similarity between two or more words, a part of speech, or other techniques. In some embodiments of the invention, the system and method may include receiving information related to a verbalization made by a user. This information may be received from a speech engine. The information may include textual information, or other information, that may represent a plurality of preliminary interpretations of the verbalization made by the user.

FIG. 1



Best Available Copy

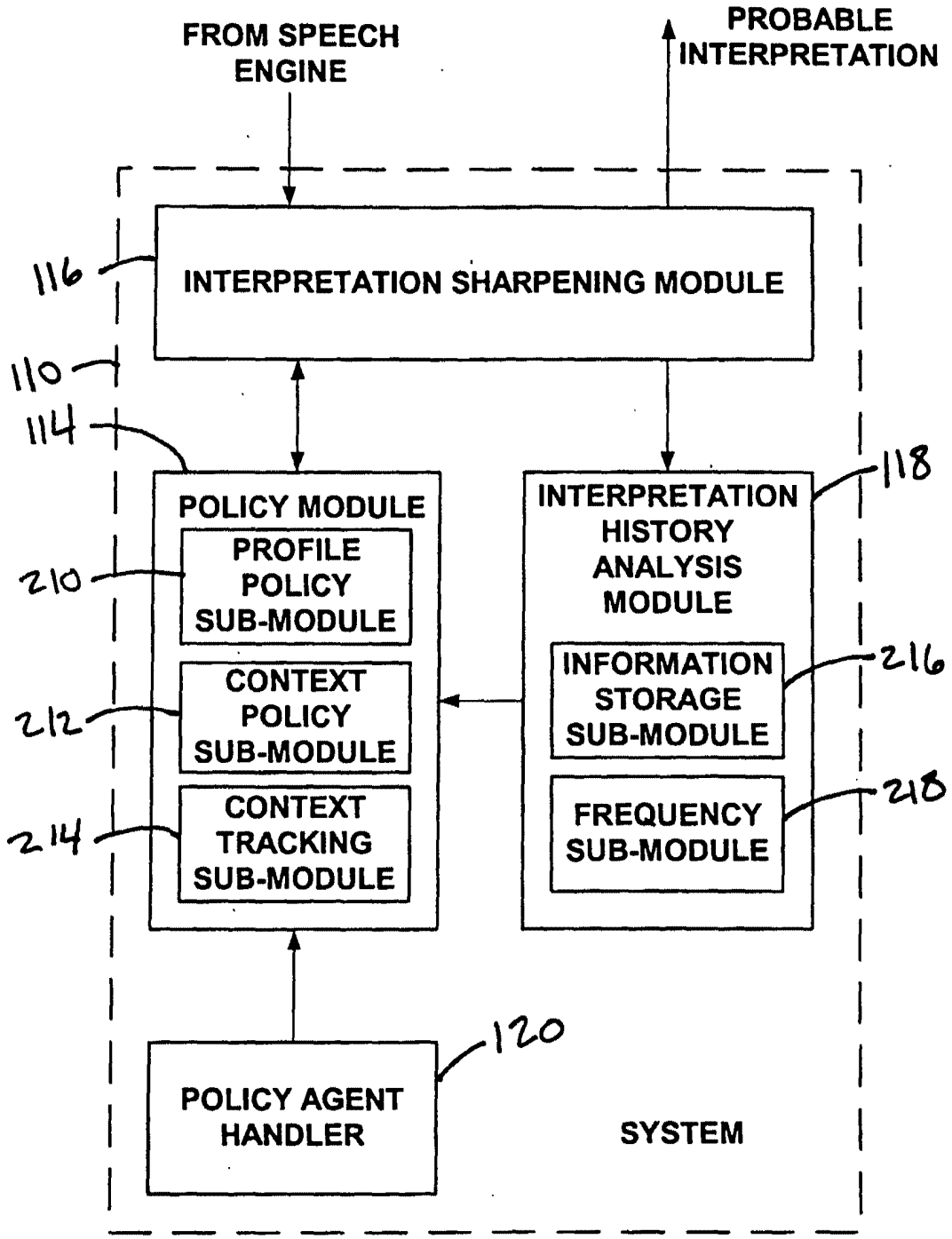
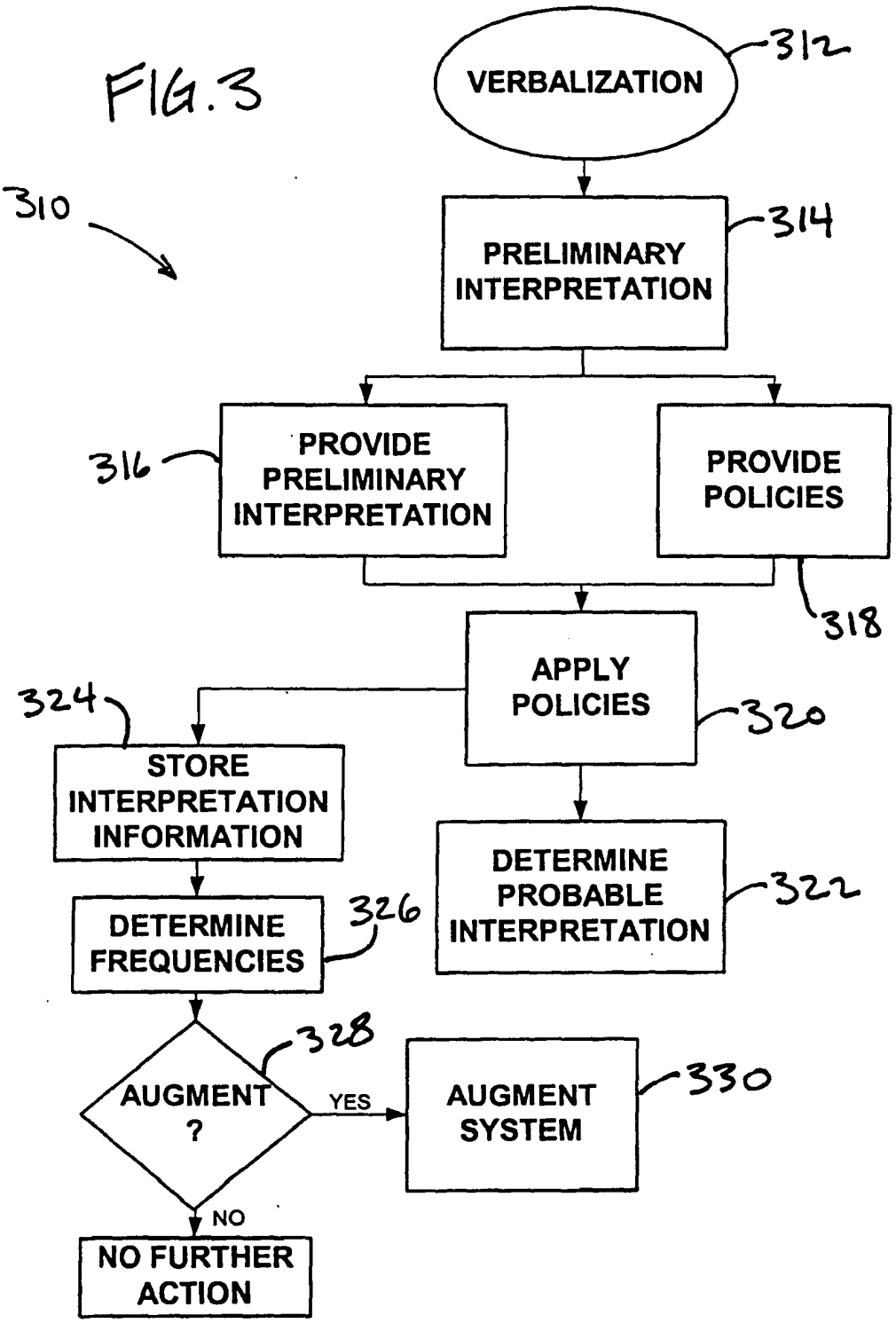


FIG. 2

Best Available Copy

FIG. 3



Best Available Copy

PATENT APPLICATION SERIAL NO _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

09/01/2005 HLE333 00000020 033975 60712412
01 FC:2005 100.00 DA

PTO-1556
(5/87)

*U.S Government Printing Office: 2002 — 466-247/69033

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: ROBERT A. KENNEWICK

Application No.: TBD

Filed: August 31, 2005

For: DYNAMIC SPEECH SHARPENING

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPLICATION DATA SHEET
37 C.F.R. § 1.76

BIBLIOGRAPHIC DATA

1. Applicant information

First applicant: ROBERT A. KENNEWICK
Citizenship: United States of America
Residence: 2204 S. Eddy Street, Seattle, Washington 98108

2. Correspondence information

Correspondence for this application should be addressed as follows:

Customer No.: 00909

3. Application information

Title of Invention: DYNAMIC SPEECH SHARPENING

Docket number assigned to this application: 090611-0312471
Suggested Classification: Class:
Subclass:
Technology Center to which subject matter is assigned:

Total number of drawing sheets: 3
Type of application: Provisional

Secrecy order under § 5.2:

This application does not disclose subject matter of an application which is under a secrecy order pursuant to § 5.2.

4. Representative information

The following have a power of attorney or authorization of agent in this application:

Customer No.: 00909

Date: August 31, 2005

Sean L. Ingram
Registration No. 48,283
PILLSBURY WINTHROP
SHAW PITTMAN LLP
P.O. Box 10500
McLean, VA 22102
(703) 905.2107
Customer No. 00909