

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO. LTD. and SAMSUNG ELECTRONICS
AMERICA, INC.,
Petitioners,

v.

VB ASSETS, LLC,
Patent Owner

IPR2025-00869

U.S. Patent No. 8,886,536

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,886,536

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

TABLE OF CONTENTS

I.	Grounds for Standing.....	1
II.	Identification of Challenge	1
	A. Prior Art.....	1
	B. Grounds For Challenge	2
III.	The '536 Patent.....	2
	A. Patent Overview	2
	B. Technical Background.....	4
	C. Level of Ordinary Skill	5
	D. Claim Construction.....	6
IV.	GROUND 1: O'Neill-Jones Combination Renders Obvious Claims 1- 12, 16-23, 30, 32, 50-55.....	7
	A. Combination Overview	7
	1. O'Neill.....	7
	2. Jones.....	8
	3. Motivation to Combine	9
	B. Independent Claims 1 and 32	11
	1. Preambles [1P]/[32P]	11
	2. Receiving Natural Language Request [1A]/[32A]	11
	3. Providing a Response [1B]/[32B].....	11
	4. Receiving a Second Natural Language Utterance [1C]/[32C].....	12
	5. Performing Speech Recognition on One or More Words in Second Natural Language Utterance [1D]/[32D].....	12

6.	Determining Domain Information For Recognized Words Based on First Natural Language Utterance [1E]/[32E].....	13
7.	Interpreting the Second Natural Language Utterance Based on the Domain Information [1F]/[32F].....	14
8.	Providing Recognized Words to First and Second Domain Agents [1G]/[32G]	15
9.	Obtaining Interpretations of Second Natural Language Utterance from First and Second Domain Agents [1H]/[32H] and [1I]/[32I]	15
10.	Determining Interpretation [1J]/[32J]	16
11.	Determining Promotional Content [1K]/[32K]	17
12.	Presenting the Promotional Content [1L]/[32L].....	17
C.	Claim 2	18
1.	Utterances Received During Interactive Session [2A]	18
2.	Determining an Activity Relating to One or More of the Natural Language Utterances [2B]	18
3.	Determining the Promotional Content Based on the Activity [2C]	20
D.	Claim 3	20
1.	Determining Context Information Based on the First Natural Language Utterance [3A].....	20
2.	Interpreting the Second Natural Language Utterance Based on Context Information [3B].....	22
E.	Claim 4	23
F.	Claim 5	24
1.	Associating the Interpretations With Interpretation Scores [5A]-[5B]	24

2.	Determining the First Interpretation Based on the First and Second Interpretation Score [5C].....	25
G.	Claim 6	25
H.	Claims 7, 51	26
1.	Receiving a Third Natural Language Utterance [7A]/[51A]	26
2.	Determining an Interpretation of the Third Natural Language Utterance Based on the Promotional Content [7B]/[51B]	27
I.	Claim 8	28
J.	Claim 9	28
K.	Claim 10	29
1.	Performing Speech Recognition on Third Natural Language Utterance [10A].....	29
2.	Determining That One or More Reference Words Refer to an Item Associated with the Promotional Content [10B]	30
3.	Initiating a Purchase Transaction [10C]	31
L.	Claim 11	32
M.	Claims 12, 52.....	32
N.	Claim 16	33
1.	Identifying One or More Requests Associated With Second Natural Language Utterance [16A].....	33
2.	Obtaining Promotional Content Related to Request [16B]	33
O.	Claim 17	34
1.	Determining One or More Activities in the One or More Requests [17A].....	34

2.	Performing the One or More Activities [17B].....	35
3.	Determining Promotional Content Based on Activity [17C]	35
P.	Claim 18	35
Q.	Claim 19	36
R.	Claim 20	37
S.	Claim 21	38
1.	Determining a Reservation For the User in Response to One or More Requests [21A].....	38
2.	Determining Promotional Content Based on Information Related to the Reservation [21B].....	38
T.	Claim 22	39
1.	Determining a First Reservation in Response to User Request [22A]	39
2.	Determining Promotional Content Based on Alternative Reservation [22B]	40
U.	Claim 23	41
V.	Claim 30	41
1.	Obtaining User Profile Information [30A]	41
2.	Identifying Categories of Items Based on Prior User Interactions [30B]	42
3.	Determining a Promotional Item Associated with One or More Categories [30C]	43
W.	Independent Claim 50	44
1.	[50P]-[50E] and [50G].....	44

2.	Determining Promotional Content Associated with One of the Categories Based on the First and Second Natural Language Utterances [50F].....	45
X.	Independent Claim 53	46
Y.	Independent Claim 54	46
1.	[54P]-[54E] and [54H]-[54I].....	46
2.	Determining Applications for Processing Requests [54F]	47
3.	Identifying Categories of Items Related to Applications Based on Prior Use Interactions [54G]	48
Z.	Independent Claim 55	48
V.	GROUND 2: O'Neill-Jones-Franco Combination Renders Obvious Claims 13-15, 24, 33-38	49
A.	Combination Overview	49
1.	Franco.....	49
2.	Motivation to Combine	50
B.	Claim 13	51
1.	[13A]-[13B]	51
2.	Processing the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [13C].....	51
C.	Claims 14 and 34.....	56
D.	Claims 15 and 35	57
1.	[15A]/[35A]	57
2.	Determining Other Promotional Content Based on the Reinterpretation [15B]/[35B].....	58

3.	Presenting the Promotional Content to the User [15C]/[35C]	59
E.	Claim 24	59
1.	Determining at Least One Request is Incomplete or Ambiguous [24A]	59
2.	Monitoring Interactions of the User with the Promotional Content [24B].....	60
3.	Interpreting the Incomplete or Ambiguous Request Based on the Interaction [24C]	61
F.	Independent Claim 33	63
G.	Independent Claim 36	64
1.	[36P]-[36D], [36F]-[36I].....	64
2.	[36E].....	64
H.	Independent Claim 37	65
1.	[37P]-[37E] and [37H]-[37I].....	65
2.	Determining Promotional Content Relating to Request [37F].....	65
I.	Independent Claim 38	66
VI.	GROUND 3: O'Neill-Jones-Wang Combination Renders Obvious Claims 25-26, 31, 39-43	67
A.	Combination Overview	67
1.	Wang	67
2.	Motivation to Combine	68
B.	Claim 25	69

1.	First Request Associated with First Application and Second Request Associated with Second Application [25A]	69
2.	Promotional Content Related to One or More of the First or Second Request [25B]	71
C.	Claims 26, 40.....	71
D.	Claim 31	71
1.	[31A]-[31B] and [31E]	71
2.	Determining Application for Processing One or More Requests [31C].....	72
3.	Identifying Categories of Items Based on Prior Use Interactions Related to One or More Applications [31D]	72
E.	Independent Claim 39	73
F.	Dependent Claim 41	74
G.	Dependent Claim 42	74
H.	Independent Claim 43	74
VII.	GROUND 4: O’Neill-Jones-Howard Combination Renders Obvious Claims 27-29, 44-49	75
A.	Combination Overview	75
1.	Howard.....	75
2.	Motivation to Combine	77
B.	Claim 27	79
1.	First Request Processed by a First Device and Second Request Processed by Second Device [27A].....	79
2.	Promotional Content Relates to One or More Requests [27B]	80

C.	Claims 28, 45.....	80
D.	Claims 29, 46.....	81
E.	Independent Claim 44	82
F.	Dependent Claim 47	82
G.	Dependent Claim 48.....	82
H.	Independent Claim 49	82
VIII.	Mandatory Notices.....	83
A.	Real Party In Interest.....	83
B.	Related Matters.....	83
C.	Notice of Counsel and Service Information.....	84
IX.	Conclusion	85

EXHIBIT LIST

Exhibit	Description
1001	U.S. Patent 8,886,536 to Freeman & Kennewick (“the ’536 patent”)
1002	Prosecution History for U.S. Patent 8,886,536
1003	Declaration of Stuart Lipoff in Support of Inter Partes Review of U.S. Patent 8,886,536
1004	Curriculum Vitae of Stuart Lipoff
1005	U.S. Patent Application Publication 2005/0278180 to O’Neill & Hanna (“O’Neill”)
1006	U.S. Patent Application Publication 2007/0174258 to Jones & Bostic (“Jones”)
1007	U.S. Patent Application Publication 2010/0125458 to Franco et al. (“Franco”)
1008	U.S. Patent 7,609,829 to Wang et al. (“Wang”)
1009	U.S. Patent 6,513,006 to Howard & Junqua (“Howard”)
1010	Memorandum Order Regarding Claim Construction from <i>VB Assets, LLC v. Amazon.com Servs. LLC</i> , No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019)
1011	Second Amended Joint Claim Construction Chart from <i>VB Assets, LLC v. Amazon.com Servs. LLC</i> , No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019)
1012	U.S. Patent 7,620,549 to Di Cristo et al. (“Di Cristo”)
1013	U.S. Patent 5,321,608 to Namba (“Namba”)
1014	Niels Ole Bernsen et al., <i>Designing Interactive Speech Systems: From First Ideas to User Testing</i> , Springer (1998) (“Bernsen”)

Exhibit	Description
1015	Xuedong Huang et al., <i>Spoken Language Processing</i> , Prentice Hall PTR (2001) (“Huang”)
1016	Michael F. McTear, <i>Spoken Dialogue Technology</i> (2004) (“McTear”)
1017	Andrew Kehler, <i>The Effect of Establishing Coherence in Ellipsis and Anaphora Resolution</i> , 31st Annual Meeting of the Association for Computational Linguistics (1993) (“Kehler”)
1018	Sebastian Möller, <i>Quality of Telephone-Based Spoken Dialogue Systems</i> , Springer (2005) (“Möller”)
1019	Phillip R. Cohen et al., <i>An Open Agent Architecture</i> , Proceedings of the 1994 AAAI Spring Symposium, Software Agents (1994) (“Cohen”)
1020	Stephanie Seneff et al., <i>Multimodal Discourse Modelling in a Multi-User Multi-Domain Environment</i> , 4th International Conference on Spoken Language Processing (Oct. 1996)
1021	Stephanie Seneff et al., <i>GALAXY-II A Reference Architecture for Conversational System Development</i> , Proceedings of the 5th International Conference on Spoken Language Processing (Nov. 1998)
1022	Ian O’Neill et al., <i>Cross Domain Dialogue Modelling: An Object-Based Approach</i> , Proceedings of the 8th International Conference on Spoken Language Processing (2004) (“O’Neill (2004)”)
1023	Botond Pakucs, <i>SesaME: A Framework for Personalised and Adaptive Speech Interfaces</i> , Proceedings of the 2003 EACL Workshop on Dialogue Systems: Interaction, Adaptation, and Styles of Management (2003) (“Pakucs”)
1024	Mikio Nakano et al., <i>Grounding New Words on the Physical World in Multi-Domain Human-Robot Dialogues</i> , Papers from the 2010 AAAI Fall Symposium (2010) (“Nakano”)

Exhibit	Description
1025	Honghua (Kathy) Dai et al., <i>Detecting Online Commercial Intention (OCI)</i> , Proceedings of the 15th International Conference on World Wide Web (2006) (“Dai”)
1026	Interactive Advertising Bureau, <i>IAB Internet Advertising Revenue Report</i> (May 2007)
1027	Joe Plummer et al., <i>The Online Advertising Playbook</i> , (2007) (“Plummer”)
1028	Ganesh Iyer et al., <i>The Targeting of Advertising</i> , INSEAD Working Paper Series (Feb. 2004) (“Iyer”)
1029	Alex Kinnier, <i>Why We’re Buying DoubleClick</i> , Google Blog (June 26, 2007), https://googleblog.blogspot.com/2007/06/why-were-buying-doubleclick.html
1030	Daniel C. Fain & Jan O. Pedersen, <i>Sponsored Search: A Brief History</i> , 32(2) Bulletin of the American Society for Information Science & Technology 12 (2006) (“Fain”)
1031	Press Release, Google, Google Launches Self-Service Advertising Program (Oct. 23, 2000), http://googlepress.blogspot.com/2000/10/google-launches-self-service.html
1032	Eric J. Sinrod, Duane Morris LLP, <i>Internet Advertising Via “Behavioral Targeting”</i> , USA Today.com (Sept. 28, 2004) (“Duane Morris”)
1033	Kate Burns, <i>Ten Golden Rules to Search Advertising</i> , 6(3) Interactive Marketing 248 (2005) (“Burns”)

Samsung Electronics Co. Ltd. and Samsung Electronics America, Inc.

(“Petitioners”) petition for *inter partes* review of claims 1-55 of U.S. Patent 8,886,536 (“the ’536 patent”; EX-1001).

I. Grounds for Standing

Petitioners certify the ’536 patent is available for IPR; Petitioners are not barred or estopped.

II. Identification of Challenge

A. Prior Art

The ’536 patent, filed on September 3, 2013, claims priority to U.S. Patent 7,818,176, filed February 6, 2007. While Petitioners do not agree that the ’536 patent is entitled to priority any earlier than September 3, 2013, each applied reference was filed or published before February 6, 2007.

1. **U.S. Publication 2005/0278180 to O’Neill and Hanna** (“O’Neill”; EX-1005), published December 15, 2005, is prior art under 35 U.S.C. §102(b).
2. **U.S. Publication 2007/0174258 to Jones and Bostic** (“Jones”; EX-1006), filed December 29, 2006, is prior art under 35 U.S.C. §102(e).
3. **U.S. Publication 2010/0125458 to Franco et al.** (“Franco”; EX-1007), filed July 13, 2006, is prior art under 35 U.S.C. §102(e).
4. **U.S. Patent 7,609,829 to Wang et al.** (“Wang”; EX-1008), filed March 17, 2004, is prior art under 35 U.S.C. §102(e).

5. **U.S. Patent 6,513,006 to Howard and Junqua** (“Howard”; EX-1009), issued January 28, 2003, is prior art under 35 U.S.C. §102(b).

B. Grounds For Challenge

Ground		Claims	Prior Art
1	103	1-12, 16-23, 30, 32, 50-55	O’Neill and Jones
2	103	13-15, 24, 33-38	O’Neill, Jones and Franco
3	103	25-26, 31, 39-43	O’Neill, Jones and Wang
4	103	27-29, 44-49	O’Neill, Jones and Howard

III. The ’536 Patent

A. Patent Overview

The ’536 patent discloses a system (depicted in Figure 1, below) for “selecting and presenting advertisements based on natural language processing of voice-based input.” (EX-1001, 1:29–31.)

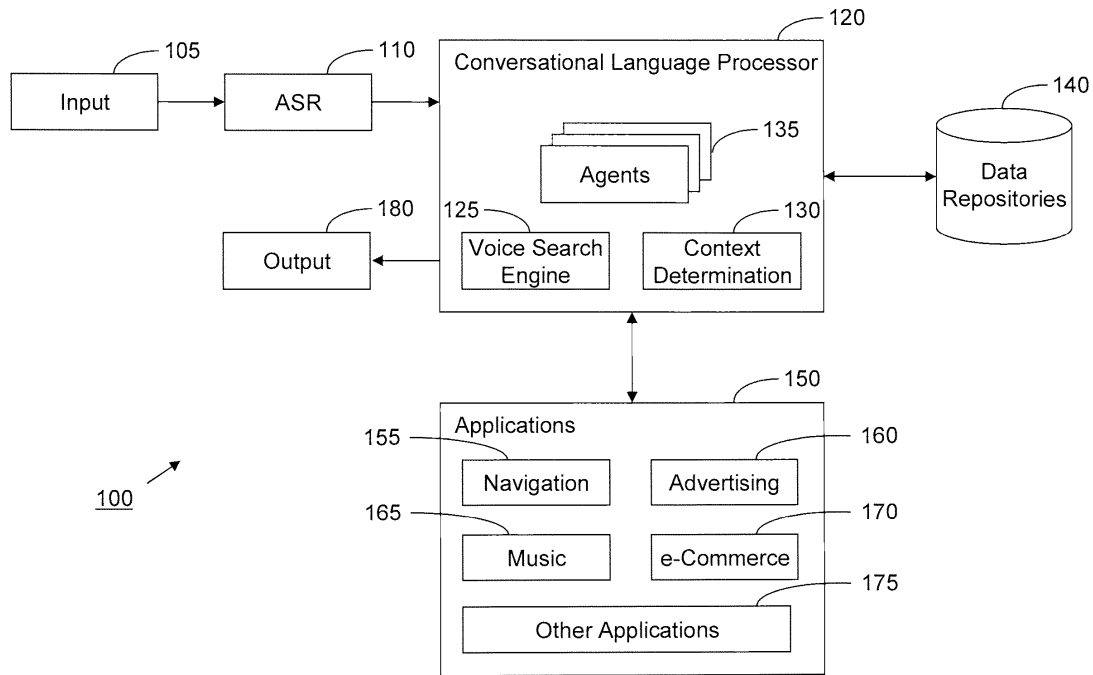


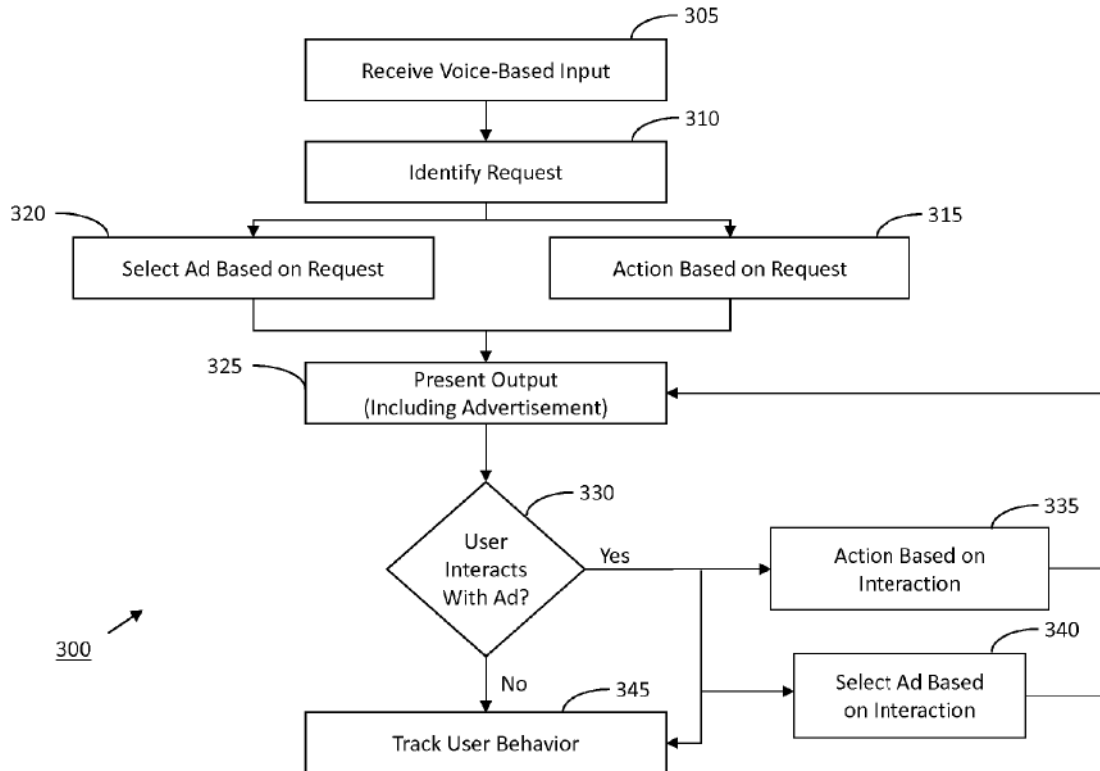
Fig. 1

'536 Patent, Figure 1

The system receives a voice-based utterance and “may enable users to ... retrieve information or perform various tasks” such as provide navigation routes or make phone calls. (EX-1001, 3:24-27; 3:52-54; 7:39-59.)

The system processes the input using an automatic speech recognizer that generates one or more preliminary interpretations of the utterance. (EX-1001, 3:59-62.) The requests are further processed by a conversational language processor that may “establish a context for a current utterance by having domain agents 135 competitively generate a context-based interpretation of the utterance.” (EX-1001, 4:50-55.)

Figure 3 illustrates an “exemplary method for selecting and presenting advertisements.” (EX-1001, 3:12-14.)



'536 Patent, Figure 3

Information in the input is also communicated to an advertising server to select one or more advertisements related to the request (step 320). (EX-1001, 8:27-33.)

B. Technical Background

Sophisticated speech dialogue systems were well known and commercially available before the priority date. Many of these systems relied on, one, more or all of the following:

- **Speech Recognizer:** Software and hardware for converting the utterance to text.

- **Semantic Analyzer:** Interprets the text to understand meaning of the utterance.
- **Dialogue Manager:** Interprets the meaning of the utterance in broader context using, for instance, dialog history.
- **Response Generator:** Creates natural language replies.
- **Speech Synthesizer:** Converts the response into speech, and transmits it back to the user.

(EX-1003, ¶44.) Further, these systems expressly incorporated context to facilitate the interpretation of meaning through the dialog manager, including the use of domain-specific knowledge to interpret meaning.

Likewise, commercially available systems had already been considering context to select appropriate ads in conjunction with speech recognition tools that were being used to access information (e.g., calling directory assistance or searching the web). Finally, by the priority date of the '536 patent, it was standard practice to use domain agents and user tracking to target users with appropriate promotional content. (EX-1003, ¶¶32-94.)

C. Level of Ordinary Skill

A person of ordinary skill in the art (“POSITA”) would have a bachelor’s degree in computer science, computer engineering, electrical engineering, or related field in computing technology, and two years of experience with automatic speech recognition and natural language understanding, or equivalent education, research experience, or knowledge. (EX-1003, ¶¶124-125.)

D. Claim Construction

In *VB Assets, LLC v. Amazon.com Servs. LLC*, No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019) (“Amazon-VBA-Litigation”) the Court construed various terms as set forth below. (EX-1010.)

Claim Term	Construction
“domain agent”	“software with domain-specific behavior and information” (EX-1010, 1-2)
“speech recognition engine” / “speech recognition”	“software or hardware that recognizes the words or phrases in the natural language utterance” ¹ (EX-1010, 2)
“context”	No construction necessary (EX-1010, 2 (rejecting Amazon’s proposal that the term should mean “subject domain(s) that applies to a user utterance”))

In *Amazon.com, Inc. et al. v. VB Assets, LLC*, IPR2020-01374 and -01377, the Board “determine[d] that no express construction of any term is necessary.” (EX-1002, 451 (IPR2020-01377), 479 (IPR2020-01374).)

¹ The court in the Amazon-VBA-Litigation adopted Amazon’s proposed construction, rejecting PO’s position that a “speech recognition engine” is “software or hardware that generates one or more preliminary interpretations of a given utterance.” (EX-1011, EX. A-4.)

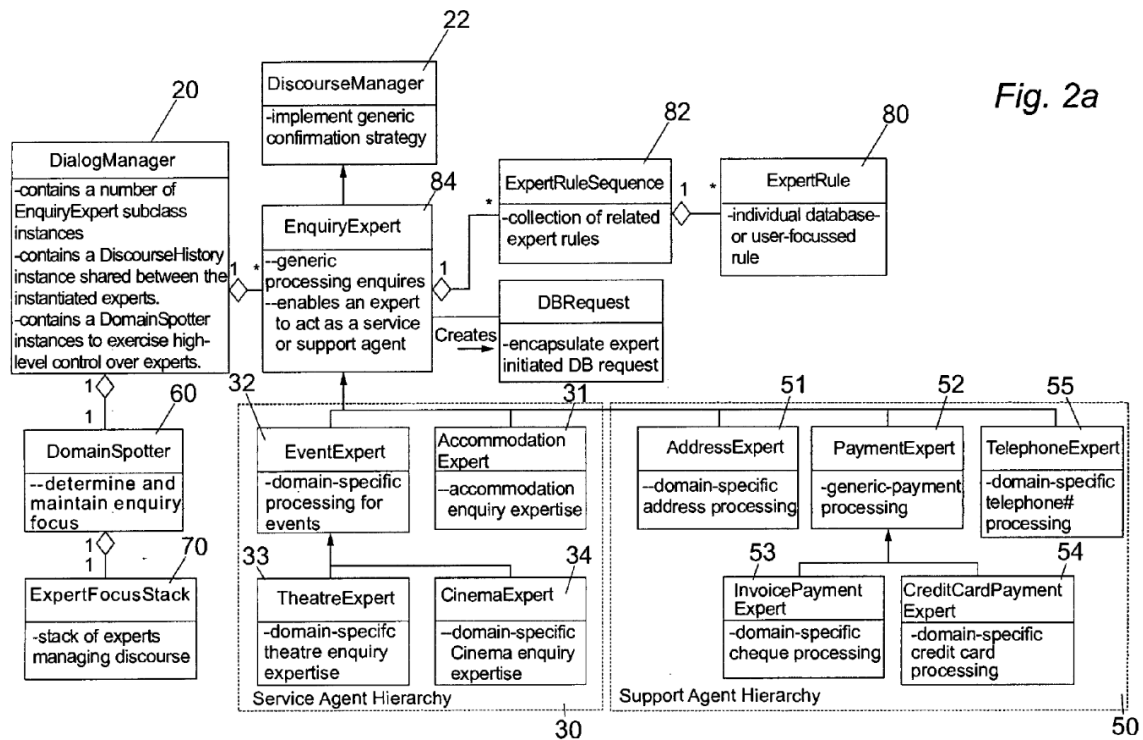
The Board need not expressly construe any term for purposes of this proceeding because Petitioners establish the challenged claims are unpatentable under the constructions entered in the Amazon-VBA-Litigation and the claims' plain meaning, where different. *See Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999).

IV. GROUND 1: O'Neill-Jones Combination Renders Obvious Claims 1-12, 16-23, 30, 32, 50-55

A. Combination Overview

1. O'Neill

O'Neill relates to "automatic dialogue system[s]" used to "assist telephone callers in completing a transaction in a well-defined business domain." (EX-1005, [0003].) To facilitate an interactive dialogue, O'Neill uses "agents" with "domain-specific functionality." (EX-1005, [0029]-[0032].) "Each agent is a specialist in a particular transactional area and uses its own domain-specific expert rules to encapsulate a skill-set for a substantial dialogue." (EX-1005, [0073].) The system (shown an exemplary schematic, below) can more efficiently assist users with various tasks, like booking a hotel. (*See, e.g.*, EX-1005, [0132]-[0154], [0264].)



O'Neill, Figure 2a

The system includes a DomainSpotter which determines the domain agent that is most appropriate to handle the user's request. (EX-1005, [0198], [0200].) The DomainSpotter picks the domain agent with the highest confidence score to handle the user's request. (EX-1005, [0202].)

2. Jones

Jones "provides targeted advertising to cellular platforms using voice recognition by identifying a keyword in a spoken phrase, looking up the keyword in an index having corresponding advertising and providing the advertising to cellular subscribers." (EX-1006, [0066].) The system leverages user information to

determine what advertisement to provide. (EX-1006, [0031].) The system further tracks user engagement with advertisements, including whether a user interacted with an advertisement or purchased an advertised product. (EX-1006, [0032]-[0033].)

The system further uses context to interpret the meaning of a user's request, such as if a user's request can have multiple meanings. (EX-1006, [0040], [0053].)

3. Motivation to Combine

A POSITA would have been motivated to combine O'Neill's teachings regarding its interactive natural language dialogue system with the Jones system for providing targeted promotional content for numerous reasons.

First, both O'Neill and Jones suggest that their systems can be combined with other speech processing systems. (EX-1006, [0025] (suggesting that the Jones system may use "conventional speech-to-text processing, also known as automatic speech recognition (ASR)"); EX-1005, [0089] ("the improved dialogue manager could also be used in other automatic dialogue systems").)

Second, the systems disclosed in O'Neill and Jones are designed to assist users with requests—like locating information about businesses—using ASR. (EX-1006, [0023] (system designed to "receive[] voice requests or queries submitted by users ... and return[] responses resulting from the processing."); EX-1005, Abstract (system designed to "receive[] a communication from a user," process the

communication, “and provide[] a response to the user.”).) As O’Neill and Jones are attempting to accomplish the same goal of receiving requests and responding to requests, a POSITA would be motivated to combine their teachings.

Third, adding the targeted advertising of Jones to O’Neill’s interactive dialogue system would monetize the system, which the system in O’Neill is already designed to handle, as O’Neill’s system includes payment processing functionality. (EX-1005, [0366]-[0370].)

Finally, the combination is nothing more than the application of a known technique (targeted advertising disclosed in Jones) to a similar known system (O’Neill’s interactive dialogue system) that operates in the same way (analyzing natural language utterances to interact with user’s requests).

A POSITA would have had a reasonable expectation of success in the combination and the results of the combination would have been predictable. Combining O’Neill with Jones would merely integrate one component (targeted advertising as disclosed in Jones) into O’Neill’s interactive dialogue system. O’Neill and Jones are also both in the same field of the ’536 patent—using speech recognition to facilitate access to information. (*See* EX-1001, 2:22-25; EX-1005, Abstract, [0023]-[0027]; EX-1006, [0009]; EX-1003, ¶¶160-169.)

B. Independent Claims 1 and 32

1. Preambles [1P]/[32P]

The O'Neill-Jones combination discloses a “*system*” that practices a “*method*” of “*providing promotional content related to one or more natural language utterances and/or responses.*” The Jones system “provides targeted advertising to cellular platforms using voice recognition.” (EX-1006, [0066]; EX-1003, ¶173.) O'Neill also discloses a system that uses “*processing units*” to engage in an interactive dialogue session with a user that includes receiving a user utterance, processing the utterance and responding. (EX-1005, [0023]-[0027]; EX-1003, ¶¶171-76.)

2. Receiving Natural Language Request [1A]/[32A]

The O'Neill-Jones combination discloses this limitation. O'Neill discloses a method which starts with “receiving a communication from the user.” (EX-1005, [0052]-[0058]; *see also* [0023]-[0027]; EX-1003, ¶178.) Jones similarly teaches that users “submit voice requests using mobile phones or other mobile networked devices” before the system selects a targeted advertisement. (EX-1006, Abstract; EX-1003, ¶181.) Both systems “*receive[]*” “*a first natural language utterance*” at “*physical processors.*” (EX-1003, ¶¶177-81.)

3. Providing a Response [1B]/[32B]

The O'Neill-Jones combination discloses this limitation. O'Neill discloses a method where “the assigned processing unit [*physical processor*] processes the

communication ... and provides a *response* to the user.” (EX-1005, [0023]-[0027]; *see also* O’Neill, [0052]-[0058]; EX-1003, ¶182-84.) Jones also discloses a system that can, in response to the user’s request, “produce automated results.” (EX-1006, [0046]; EX-1003, ¶¶182-85.)

4. Receiving a Second Natural Language Utterance [1C]/[32C]

The O’Neill-Jones combination discloses this limitation. O’Neill describes an interactive dialogue with a user that includes multiple turns of dialogue, including “*a second natural language utterance*.” (EX-1005, [0052]-[0058] (“repeating [dialogue] steps until the user fails to issue further communications or indicates that the dialogue is terminated.”); EX-1003, ¶187.) Such multiple turns of dialogue evidence “*a second natural language utterance relating to the first natural language utterance*.” (EX-1005, [0291]-[0321] (providing example dialogue between system and user related to reserving a hotel room); EX-1003, ¶186-89.)

5. Performing Speech Recognition on One or More Words in Second Natural Language Utterance [1D]/[32D]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶190-96.)

As discussed above, in the Amazon-VBA-Litigation the term “speech recognition” was construed to mean “software or hardware that recognizes the words or phrases in the natural language utterance.”

The O'Neill system includes a speech recognizer and natural language parser that recognizes one or more words in the user utterance and provides the semantically tagged phrases to the system's dialogue manager that is responsible for the interactive dialogue with the user. (EX-1005, [0013], Fig. 1, [0096] (noting that the O'Neill system is integrated into the DARPA communicator described in [0013]), [0278]; EX-1003, ¶194.) The system performs this operation for all user utterances in the interactive dialogue, including the "*second natural language utterance*." (EX-1005, [0052]-[0058]; EX-1003, ¶195.) Jones also discloses performing "speech to text conversion" using "conventional speech-to-text processing, also known as automatic speech recognition (ASR)." (EX-1006, [0025].)

6. Determining Domain Information For Recognized Words Based on First Natural Language Utterance [1E]/[32E]

The O'Neill-Jones combination discloses this limitation.

O'Neill discloses that the user's utterance is processed by the speech recognizer and natural language parser into a semantic parse of recognized words that is input to the DomainSpotter and service agents, which further processes the user utterance. (EX-1005, [0013]; [0200]-[0202], EX-1003, ¶¶198-199.) "The first step in starting a dialogue session is to identify the most appropriate agent to process a user's enquiry," which "is typically performed automatically by a DomainSpotter" determining the domain agent that can best handle the utterance.

(EX-1005, [0198], [0200].) O'Neill describes that "key words from the user's utterance like four-star and hotel fall within the accommodation domain." (EX-1005, [0294]; EX-1003, ¶202.) The system processes each user utterance according to the disclosed method, which includes "assigning the communication to one of a plurality of processing units, each of which comprises one or more processing rules, in accordance with the semantic content of the communication." (EX-1005, [0052]-[0058].) Therefore, O'Neill is continuously "*determining... domain information*" related to the user's most recent response, including the "*first natural language utterance*." (EX-1003, ¶¶197-204.)

7. Interpreting the Second Natural Language Utterance Based on the Domain Information [1F]/[32F]

The O'Neill-Jones combination discloses this limitation.

As discussed below, the O'Neill system performs a number of processing steps to "*determine an interpretation of the second natural language utterance*." To interpret the utterance, the O'Neill system "assign[s] the communication to one of a plurality of processing units, each of which comprises one or more processing rules, in accordance with the semantic content of the communication." (EX-1005, [0052]-[0058].) O'Neill's "processing units" are "*physical processors*" that use "*domain information*" in their "processing rules" to determine an interpretation of the user's utterances, including the "*second natural language utterance*." (EX-1003, ¶205.).

8. Providing Recognized Words to First and Second Domain Agents [1G]/[32G]

The O'Neill-Jones combination discloses this limitation.

In the Amazon-VBA-Litigation the term “domain agent” was construed to mean “software with domain-specific behavior and information.”

After the O'Neill system semantically parses the user's utterance, the DomainSpotter provides each domain agent “the output of the semantic parse” and those domain agents score their ability to handle responding to the utterance. (EX-1005, [0200]-[0201]; EX-1003, ¶208.) In one example, the semantic parse provides the terms “four star” and “hotel” to the domain agents. The domain agents score their individual ability to handle the response. Here, the hotel accommodation domain agent is deemed most appropriate. (EX-1005, [0294]; EX-1003, ¶¶209-11.) O'Neill thus discloses “determining which agent should handle the user's further utterances,” including the “*second natural language utterance*.” (EX-1005, [0208]; EX-1003, ¶¶206-211.)

9. Obtaining Interpretations of Second Natural Language Utterance from First and Second Domain Agents [1H]/[32H] and [1I]/[32I]

The O'Neill-Jones combination discloses this limitation.

The DomainSpotter provides the semantic parse of the user's natural language utterance to the domain agents which interpret and score the utterance to determine if it is the appropriate domain agent. (EX-1005, [0200]-[0202]; EX-1003, ¶213.)

For the first and second domain agent to determine if they are the appropriate domain agent for the user's request, the domain agents must obtain interpretations of the "*second natural language utterance*" and provide the DomainSpotter with a numerical score quantifying its ability to handle the request. (EX-1005, [0200]-[0202]; EX-1003, ¶214.) In one example, after the user asks about hotels in Dublin (a first natural language utterance), the system responds "So, you're staying in Dublin." (EX-1005, [0213].) In response, the user states (a "*second natural language utterance*"), "Yes," which the system then interprets as confirming the user is searching for a hotel in Dublin. (EX-1005, [0213]; *see also* [0261]-[0267] (transferring between domain agents), [0271] (polling domain agents to interpret request); EX-1003, ¶212-16.)

10. Determining Interpretation [1J]/[32J]

The O'Neill-Jones combination discloses this limitation.

The domain agent with the highest score is assigned by the DomainSpotter to handle the user's request, meaning that the system is "*determining the interpretation based*" on the interpretation of the domain agent with the highest score. (EX-1005, [0200]-[0202]; EX-1003, ¶218.) In other words, the O'Neill system determines the interpretation by selecting the domain agent with the highest score. (EX-1005, [0200]-[0202]; EX-1003, ¶218.) The system then assigns that domain agent to handle the user's request and continue the dialogue with the user.

(EX-1005, [0200]-[0202]; *see also* [0261]-[0267] (describing transferring between domain agents), [0271] (describing polling other domain agents to interpret request); EX-1003, ¶¶217-223.)

11. Determining Promotional Content [1K]/[32K]

The O'Neill-Jones combination discloses “*determining, by the one or more physical processors, promotional content based on the interpretation.*”

Jones discloses a system that provides “targeted advertising ... using voice recognition by identifying a keyword in a spoken phrase, looking up the keyword in an index having corresponding advertising and providing the advertising.” (EX-1006, [0066].) The alleged invention exists to “deliver targeted advertisements and track advertisement interactions in voice recognition contexts.” (EX-1001, Abstract.) A POSITA would therefore understand that the “advertising” discussed in Jones is “*promotional content.*” (EX-1003, ¶¶224-26.)

12. Presenting the Promotional Content [1L]/[32L]

The O'Neill-Jones combination discloses this limitation.

Jones discloses “using voice recognition by identifying a keyword in a spoken phrase, looking up the keyword in an index having corresponding advertising and **providing the advertising.**” (EX-1006, [0066]; *see also* EX-1006, [0041]; EX-1003, ¶228.) While O'Neill does not present promotional content to

the user, the system does present the user with information in response to user utterances via a response. (EX-1005, [0052]-[0058]; EX-1003, ¶¶227-29.)

C. Claim 2

1. Utterances Received During Interactive Session [2A]

The O'Neill-Jones combination discloses this limitation.

As discussed above (§IV.A.1), the system in O'Neill is designed to be an “*interactive session*” that includes multiple turns of dialogue (including a “*first natural language utterance*” and “*second natural language utterance*”) and where “the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically.” (EX-1005, [0077].) The O'Neill system is designed to have an “*interactive session between the user and the computer system*” with a “conversational style” rather than overarching dialogue plan or agenda. (EX-1005, [0077], [0110]; EX-1003, ¶¶230-32.)

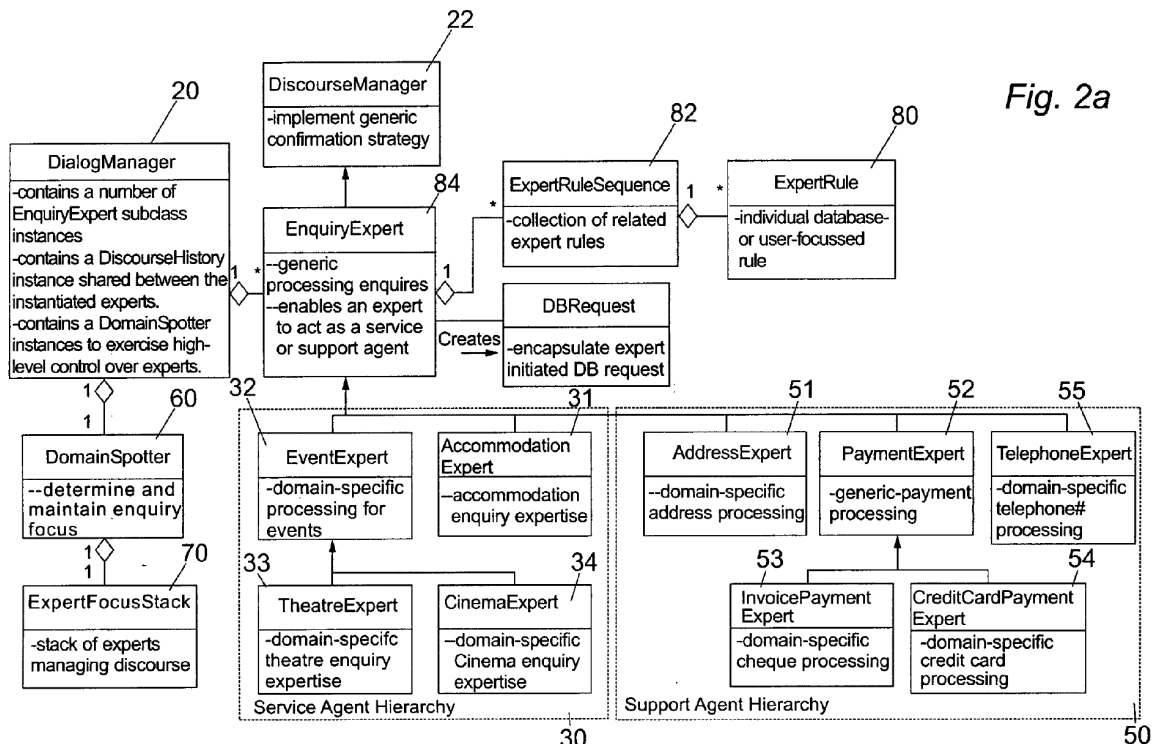
2. Determining an Activity Relating to One or More of the Natural Language Utterances [2B]

The O'Neill-Jones combination discloses this limitation.

The '536 patent specification only uses the term “activity” once:

“Advertisements may be selected based on various criteria, including ... an activity related to the input (e.g., a relation to a request, a requested application, etc.).” (EX-1001, 2:31-36.)

As discussed above, and as shown in Fig. 2a, (below) O'Neill's DomainSpotter determines whether the user's utterances relate to certain types of "activit[ies] relating to one or more of" the first or second utterance, such as reserving a hotel room or booking cinema tickets, before a specific domain agent is assigned. (EX-1005, [0200]-[0204], [0394]-[0405]; EX-1003, ¶235.)



O'Neill, Figure 2a

Jones similarly discloses determining "an activity relating to one or more of" the first or second utterance, such as determining that a user's request related to the activity of obtaining restaurant information. (EX-1006, [0031]; EX-1003, ¶233-37.)

3. Determining the Promotional Content Based on the Activity [2C]

The O’Neill-Jones combination discloses this limitation.

Jones discloses a system that determines promotional content using a number of factors to weight and determine the appropriate promotional content, including “*activity*.” (EX-1006, [0029].) In one example, the system determines the promotional content—advertisements related to Broadway shows—after the user indicated they were interested in attending a performance at The Kennedy Center for the Performing Arts. (EX-1006, [0029]; *see also* EX-1006, [0031] (example based on restaurants in Indiana); EX-1003, ¶239.)

While O’Neill does not determine promotional content, O’Neill does disclose using all of the user’s utterances and the system’s responses to process future user utterances. (EX-1005, [0103]-[0104] (O’Neill contains “components for storing information acquired or generated during a dialogue session” and used by the system to “prepare system utterances in accordance with their generic and domain specific rules.”); EX-1003, ¶242.) Therefore, the O’Neill-Jones combination would use all information when “*determining the promotional content*.” (EX-1003, ¶¶238-42.)

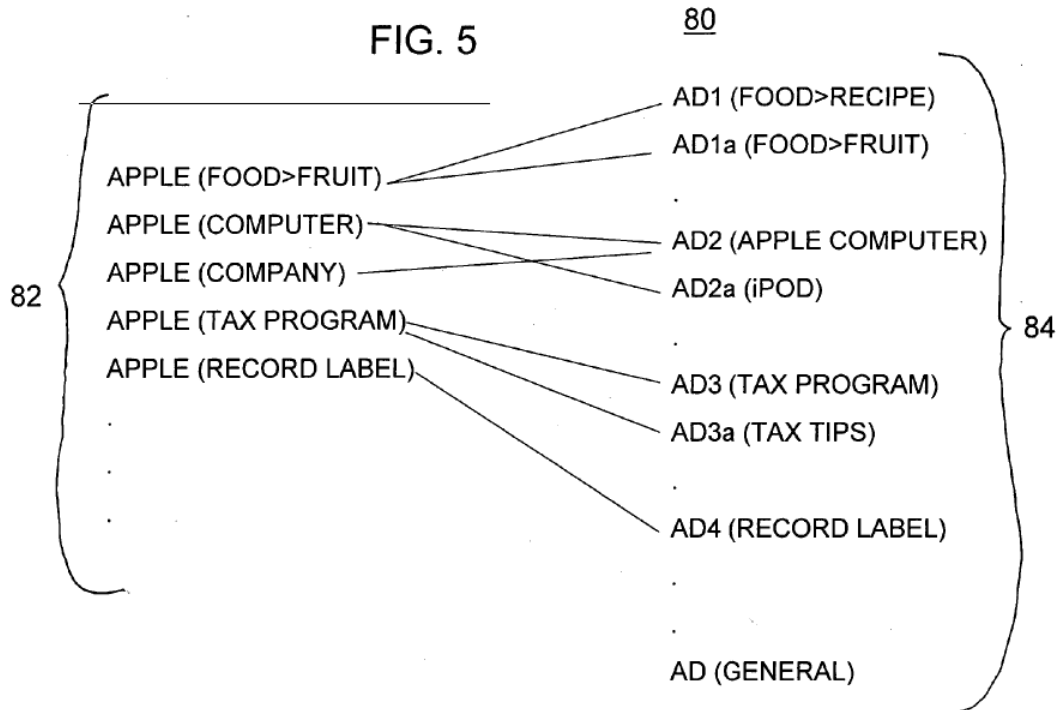
D. Claim 3

1. Determining Context Information Based on the First Natural Language Utterance [3A]

The O’Neill-Jones combination discloses this limitation.

The '536 patent discloses it “establish[es] a context for a current utterance by having domain agents 135 competitively generate a context-based interpretation of the utterance.” (EX-1001, 4:50-55.) O’Neill discloses a substantively identical process for “*determining... context*” using domain agents to competitively generate a context-based interpretation of the utterance by scoring alternative interpretations and selecting a highest scoring interpretation as discussed above in §IV.B. (EX-1003, ¶¶247-48.)

Jones also determines “*context information*” for recognized words. (EX-1006, [0037].) For example, the system associates “Reggie Miller” to the context of the “Indiana Pacers.” (EX-1006, [0037].) Similarly, Jones (as shown in Figure 5, below) discloses establishing context to select advertisements, particularly for words that have multiple meanings such as “apple.” (EX-1006, [0061].)



Jones, Figure 5

Moreover, the '536 patent admits “context-based interpretations and responses to a voice-based input” was already known in the art and not a novel feature of the alleged invention. (EX-1001, 4:59-5:3; EX-1003, ¶¶243-52.)

2. Interpreting the Second Natural Language Utterance Based on Context Information [3B]

The O'Neill-Jones combination discloses “*wherein the interpretation is determined further based on the context information.*”

As discussed above, to process the user's utterance, the domain agent uses domain-specific knowledge and “processing rules” to interpret the user's utterance. (EX-1005, [0021]-[0030], claim 26.) The “processing rules” leverage the context established by the system for the user's utterance. (EX-1003, ¶254.) For example,

following the user's utterance containing the words "four star" and "hotel," the system interprets the user's utterance in the context of booking a hotel reservation and uses the domain specific "processing rules" for booking a hotel reservation. (EX-1005, [0291]-[0321]; EX-1003, ¶256.)

O'Neill also "components for storing information acquired or generated during a dialogue session," which forms the system's "discourse knowledge," and is used by the system to "prepare system utterances in accordance with their generic and domain specific rules." (EX-1005, [0103]-[0104]; EX-1003, ¶257.) The system can then "retrieve information collected at an earlier point in a dialogue to assist in a current query," e.g., interpreting a "*second natural language utterance*." (EX-1005, [0075]-[0076]; EX-1003, ¶257.)

Similarly, Jones uses context information to interpret words with multiple meanings like "apple" in the example above discussed in §IV.D.1. (EX-1003, ¶253-58.)

E. Claim 4

The O'Neill-Jones combination discloses this limitation.

As discussed above in §IV.B.8-10, in the O'Neill system each of the "*first domain agent*" and the "*second domain agent*" determine an interpretation of the user's utterances by scoring the parse of the natural language utterance "against the semantic categories that it can process." (EX-1005, [0201]-[0202]; *see also* [0261]-

[0267] (describing transferring between domain agents), [0271] (describing polling other domain agents to interpret request).) In other words, each of O'Neill's domain agents determines an interpretation of the utterance based on its specific domain information and processing rules to score its ability to handle the user's request. (EX-1005, [0201]-[0202]; EX-1003, ¶¶259-60.)

F. Claim 5

1. Associating the Interpretations With Interpretation Scores [5A]-[5B]

The O'Neill-Jones combination discloses this limitation.

As discussed above, O'Neill discloses that each of the domain agents receives and processes the second natural language utterance (which it receives from the natural language parser), provides a "*first interpretation*" and "*second interpretation*," and provides a "*score*" for each interpretation by scoring it "against the semantic categories that it can process." (EX-1005, [0201]-[0202]; EX-1003, ¶264.) In one example, the utterance "I'd like to go see THE LION KING 2" is processed by two domain agents, the EventExpert and CinemaExpert. Both domain agents interpret the user's utterance and score with its interpretation. The CinemaExpert scored higher because the utterance is interpreted as a request related to movies and assigned to the CinemaExpert domain agent. (EX-1005, [0203]-[0204]; EX-1003, ¶¶262-65.)

2. Determining the First Interpretation Based on the First and Second Interpretation Score [5C]

The O'Neill-Jones combination discloses this limitation.

O'Neill discloses that the domain agent “that attains the highest score is deemed to be the most appropriate handling agent for the user’s enquiry” and is “selected to apply its domain-specific heuristics to the more detailed processing of the user’s enquiry.” (EX-1005, [0201]-[0202]; EX-1003, ¶267.) In the Lion King 2 example above, the CinemaExpert domain agent scored higher than the EventExpert and therefore the CinemaExpert’s interpretation of the user’s utterance determined the system’s interpretation. (EX-1005, [0203]-[0204].) In another example, “if the user had been asking about hotels in Belfast” the domain agent related to hotel accommodations “might score highest and so become the handling agent.” (EX-1005, [0201]-[0202].) O'Neill therefore includes “*determining the first interpretation*” “*based on the first interpretation score and the second interpretation score.*” (EX-1003, ¶¶266-69.)

G. Claim 6

The O'Neill-Jones combination discloses this limitation.

As discussed above in §IV.F.2, in the O'Neill system, the scores returned by each domain agent are compared and the highest score is selected to handle the user’s request. (EX-1005, [0201]-[0202]; EX-1003, ¶271.) The interpretation is

therefore determined “*based on the first interpretation score being greater than the second interpretation score.*” (EX-1003, ¶¶270-71.)

H. Claims 7, 51

1. Receiving a Third Natural Language Utterance [7A]/[51A]

The O’Neill-Jones combination discloses this limitation.

As discussed above, O’Neill discloses a system that engages in an interactive “dialogue” between a user and the system, which is an “interactive session” where multiple rounds of communication between the user and system can occur, including a “*third natural language utterance.*” (EX-1005, [0052]-[0058]; EX-1003, ¶274; *see also* EX-1005, [0023]-[0027].) In one example, O’Neill discloses:

System: Welcome to the Communicator. How can I help you?

User [*first utterance*]: I’d like to book a four-star hotel in Belfast [...].

System: So that’s an accommodation booking for a four-star hotel in Belfast [...]. Please wait while I check the database...

System: You can choose between the following hotels: the Europa, the Malone Lodge, the Stormont or the Holiday Inn

User [*second utterance*] : Could you make that a three-star?

System: OK, got you. That’s three-star accommodation. Is that OK?

User [*third utterance*]: Yes, that’s right.

(EX-1005, [0291]-[0320]; EX-1003, ¶276; *see also* [0287].) In this example, the third utterance “*relate[s]*” to both the first and second utterance. (EX-1003, ¶¶273-77.)

2. Determining an Interpretation of the Third Natural Language Utterance Based on the Promotional Content [7B]/[51B]

The O’Neill-Jones combination discloses this limitation.

While O’Neill does not explicitly disclose the presentation of promotional content, the O’Neill system uses all information from the interactive dialogue session, including the “*third natural language utterance*” to interpret and respond to the user’s utterances, including the promotional content when combined with Jones, as evidenced by the example dialogue above in §IV.H.1. As discussed above, the O’Neill-Jones combination would use this dialogue as an opportunity to provide promotional content, for example, after the user asks for a three star hotel. The user’s third utterance (“Yes, that’s right”) would be interpreted in the context of the system’s prior response to the user and the system would “*determin[e] ... an interpretation of the third natural language utterance based on the promotional content.*”

O’Neill’s system contains “components for storing information acquired **or generated** during a dialogue session,” which in a combined O’Neill-Jones system would include promotional content that was generated and presented to the user. (*See* EX-1005, [0103].) This data forms the system’s “discourse knowledge,”

which is used by the system to “prepare system utterances in accordance with their generic and domain specific rules.” (EX-1005, [0103]-[0104]; EX-1003, ¶283.)

The system can then “retrieve information collected at an earlier point in a dialogue to assist in a current query,” meaning that prior information generated by the system during a dialogue session—like promotional content disclosed in Jones—is used to assist in interpretation and processing of future user utterances, such as the “*third natural language utterance*.” (See EX-1005, [0075]-[0076]; EX-1003, ¶¶278-86.)

I. Claim 8

The O’Neill-Jones combination discloses this limitation.

As discussed above in §IV.C, the O’Neill system is designed to be an interactive session that includes multiple turns of dialogue and where “the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically.” (EX-1005, [0077].) The O’Neill system is designed to have an interactive “conversational style” rather than overarching dialogue plan or agenda. (EX-1005, [0075], [0110]; EX-1003, ¶288.) It thus discloses including the first, second and third utterance in an “*interactive session between the user and the computer system*.” (EX-1003, ¶¶287-89.)

J. Claim 9

The O’Neill-Jones combination discloses this limitation.

As discussed above, O'Neill provides examples where the system receives a "*third natural language utterance*" after the "*second natural language utterance*." In the example discussed above in §IV.H.1 regarding the booking of a hotel reservation, a user makes a "*third natural language utterance*" in which the user selects a hotel and asks the system to book a double room, which occurs "*after the second natural language utterance*" in which the user requests a three star hotel. (EX-1005, [0282]; *see also* EX-1005, [0291]-[0321]; EX-1003, ¶¶290-92.)

K. Claim 10

1. Performing Speech Recognition on Third Natural Language Utterance [10A]

The O'Neill-Jones combination discloses this limitation.

As disclosed above in §IV.B.4, O'Neill's system performs speech recognition with a speech recognizer and natural language parser on all user utterances as part of the interactive dialogue session. (EX-1005, [0013], Fig. 1, [0096], [0278], [0052]-[0058]; EX-1003, ¶294.) O'Neill's system also is designed to "*recognize ... one or more reference words,*" such as the reference words "that," "three-star," "four-star" and "hotel" that the system associates with booking a hotel reservation. (EX-1005, [140]; *see also* [0291]-[0320]; EX-1003, ¶297.) In the example discussed above in §IV.H.1, the user confirms that she wants a three-star hotel in the "*third natural language utterance*." (EX-1005, [0287]; *see also* [0291]-[0321].) A POSITA would understand the system performed speech

recognition to recognize one or more words in the utterance, including reference words that refer to an item associated with promotional content, as discussed below. (EX-1003, ¶297.) For example, the word “that” may be a reference word recognized by the system as referencing the three-star hotel in the prior dialogue. (EX-1003, ¶297.)

Also as discussed above, Jones similarly discloses a system that provides “targeted advertising” “using **voice recognition by identifying a keyword in a spoken phrase**, looking up the keyword in an index having corresponding advertising and providing the advertising.” (EX-1006, [0066]; EX-1003, ¶¶293-99.)

2. Determining That One or More Reference Words Refer to an Item Associated with the Promotional Content [10B]

The O’Neill-Jones combination discloses this limitation.

The ’536 patent specification does not contain the word “item.” However, claim 11 states that one type of item is a “purchasable good” and claim 20 states that another type of item is a “media item.”

As discussed above, Jones discloses a system that provides “targeted advertising” “using voice recognition by identifying a keyword in a spoken phrase, looking up the keyword in an index having corresponding advertising and providing the advertising.” (EX-1006, [0066]; *see also* EX-1006, Abstract; EX-1003, ¶302.) For example, Jones discloses obtaining and presenting promotional

content related to iPods when the user utters “*reference words*” about Apple products and advertisements related to apple fruit when the user requests information related to apple pie recipes—iPods and apple fruit each being “*an item associated with the promotional content.*” (EX-1006, [0040].)

As discussed above, the O’Neill system also determines that one or more reference words refer to an item. The system determines that the user is referring to purchasing a movie ticket, an “*item*” after referencing Lion King 2. Therefore, a POSITA would combine these teachings in O’Neill with the targeted advertisement teachings in Jones to further enhance the selection of promotional content. (EX-1003, ¶¶300-306.)

3. Initiating a Purchase Transaction [10C]

The O’Neill-Jones combination discloses this limitation.

Jones discloses a system can initiate a “*purchase transaction associated with [an] item associated with the promotional content*” as the Jones system presents targeted advertising to users who can then “elect to purchase the full-length product [related to the advertisement], which is then supplied to him,” such as electronic merchandise like videos or MP3 audio. (EX-1006, [0055]; EX-1003, ¶313.) Such items were identified “*based on the determination that the one or more reference words refer to the item.*”

Similarly, the O’Neill system detects “[k]eywords in the user’s utterance” that indicate the system should use a payment specific domain agent to process a purchase transaction for the user. (EX-1005, [0254]; EX-1003, ¶308.) For example, the system in O’Neill can initiate a “*purchase transaction*” associated with the purchase of a cinema ticket—the “*item*”—at the user’s request. (EX-1005, [0366]-[0389]; EX-1003, ¶¶307-11.)

L. Claim 11

The O’Neill-Jones combination discloses this limitation.

As discussed above (*see* §IV.K), “*purchase transactions*” in the O’Neill-Jones combination can include videos, MP3 audio and cinema tickets, each of which is a “*purchasable good*.” (EX-1003, ¶¶312-14.)

M. Claims 12, 52

The O’Neill-Jones combination discloses this limitation

As discussed above (*see* §IV.H), in the exemplary “*third natural language utterance*” in the combination, a user seeks a three star hotel in Belfast and an ad is provided. Both O’Neill and Jones disclose systems that can initiate a purchase transaction. (*See* §IV.K.) For example, after a user selects a hotel in O’Neill, the the PaymentExpert domain agent collects the relevant purchase transaction details from the user. (EX-1005, [0230]-[0233], [0399]-[0410].) In addition, Jones discloses that those purchase transactions can be associated with the promotional

content provided to the user. (EX-1006, [0028]; EX-1003, ¶318.) For example, the Jones system tracks “if the user views an advertisement, or clicks on a link, or purchases a product or orders a service offered (sometimes referred to as ‘conversion’) related to an advertisement while awaiting results of a search.” (EX-1006, [0028].) The O’Neill-Jones combination this discloses “*initiating ... a purchase transaction associated with the promotional content*” based on an interpretation of the third utterance. (EX-1003, ¶¶315-24.)

N. Claim 16

1. Identifying One or More Requests Associated With Second Natural Language Utterance [16A]

The O’Neill-Jones combination discloses this limitation.

As discussed above in §IV.C.2, O’Neill discloses identifying requests associated with user utterances and assisting the user with those requests, such as requests to book hotel rooms or cinema tickets. (EX-1005, Fig. 2a; EX-1003, ¶¶327-328.) O’Neill provides at least one example where the system identifies a “request associated with [a] second natural language utterance,” namely, the user requests that the system search for a three star hotel in Belfast. (EX-1005, [0291]-[0321]; EX-1003, ¶¶325-29.)

2. Obtaining Promotional Content Related to Request [16B]

The O’Neill-Jones combination discloses this limitation.

As discussed above in §IV.B.4, Jones discloses providing “*promotional content*” including “*based on a determination that the promotional content relates*” to a request. For example, the Jones system obtains and presents promotional content related to iPods when the user requests information about Apple products and advertisements related to apple fruit when the user requests information related to apple pie recipes. (EX-1006, [0040]; *see also* EX-1006, [0031] (discussing request for “restaurants in Indiana”); EX-1003, ¶¶330-33.)

O. Claim 17

1. Determining One or More Activities in the One or More Requests [17A]

The O’Neill-Jones combination discloses this limitation.

As discussed above in §§IV.C; IV.N, the O’Neill system determines “*one or more activities in the one or more requests.*” (EX-1005, [0274]-[0342].) For example, if the user is trying to book a hotel reservation, she is instructing the O’Neill system to assist in that activity. (EX-1005, [0274]-[0342].) O’Neill also discloses determining other activities in the one or more requests, such as purchasing cinema or theatre tickets. (EX-1005, [0391]-[0403].)

Jones similarly discloses determining “*one or more activities in the one or more requests,*” such as determining that a user’s request relates to the activity of obtaining information regarding certain restaurants the user is interested in. (EX-1006, [0031]; EX-1003, ¶¶334-36.)

2. Performing the One or More Activities [17B]

The O'Neill-Jones combination discloses this limitation.

As discussed above in §IV.C.2, the O'Neill system “*performs*” the “*one or more activities*” for the user (searching for and booking hotel reservation or cinema ticket), after it determines what activity the user is requesting. (EX-1005, [0391]-[0403]; EX-1003, ¶338.) The system interacts with the user to search for the user and then assists the user with payment by collecting payment information from the user and other necessary information, like the user’s billing address. (EX-1005, [0366]-[0370]; EX-1003, ¶¶338.)

Jones similarly discloses a system that “*performs*” “*one or more activities*” associated with user requests, namely, obtaining information in response to user requests, such as delivering the user requested information related to a restaurant. (EX-1006, [0031]; EX-1003, ¶¶337-40.)

3. Determining Promotional Content Based on Activity [17C]

The elements in limitation [17C] are substantively identical to those in limitation [2C]. For the same reasons discussed above in §VI.C.3, the O'Neill-Jones combination discloses this limitation. (EX-1003, ¶¶341-43.)

P. Claim 18

The O'Neill-Jones combination discloses this limitation.

Jones discloses that a user can interact with the system to obtain information about a business, including the phone number. (EX-1006, [0043].) The system also

has the ability to “*determining a call to be initiated for the user*” by looking up a phone number and “*initiating*” said call in response to the request. (EX-1006, [0062]; EX-1003, ¶¶345-346.) Jones also discloses “*determining the promotional content based on the call*” using the user’s request for information about a business, including the phone number. (EX-1006, [0028]-[0030]; EX-1003, ¶¶344-47.)

Q. Claim 19

The O’Neill-Jones combination discloses this limitation.

O’Neill discloses that accommodation and detailed event booking is just one example of how the system can be used and that the “improved dialogue manager could also be used in other automatic dialogue systems such as those used for verbal control of a vehicle or route-guidance systems,” which provide a “navigation route.” (EX-1005, [0089].) A POSITA would understand that “route-guidance systems” disclosed in O’Neill are designed to “*calculat[e] the navigation route for the user.*” (EX-1003, ¶349.)

As discussed above in §IV.H.2, O’Neill-Jones combination would use all of the user’s utterances to determine the system’s responses to the user and to determine the promotional content to provide the user, including utterances related to navigation routes. (EX-1003, ¶¶350-351.)

Jones discloses using a locations associated with the user's request to determine advertisements. (EX-1006, [0030] (presenting advertisements for restaurants in a particular location after user requested information about a restaurant in that area).) Jones also discloses using the user's GPS coordinates or other geographic information to determine what promotional content to provide the user. (EX-1006, [0031] ("geographic data such as GPS coordinates may also be used to select from restaurant advertisements in Indiana to provide one in or near the designated geographic location.")) A POSITA could easily use Jones' teachings regarding using a user's geographic location to "*determinin[e] the promotional content based on the navigation route.*" (EX-1003, ¶¶348-352.)

R. Claim 20

The O'Neill-Jones combination discloses this limitation.

Jones discloses "*determining a media item to be presented in response to the one or more requests*" as in some examples the system determines that "electronic merchandise" like "MP3 downloads" may be delivered to a user in response to a user's request. (EX-1006, [0055]; EX-1003, ¶335.) The system also discloses providing an "advertisement" – which is "*promotional content*" – along with the delivery of the "*media item.*" (EX-1006, [0055].) Because Jones does not limit the type of advertisement to present to the user, a POSITA would understand that the "*promotional content [may be] different than the media item.*" (EX-1003, ¶356.)

O'Neill discloses that accommodation and event booking is one example of how the system can be used and that the "improved dialogue manager could also be used in other automatic dialogue systems." (EX-1005, [0089].) A POSITA would understand that another automatic dialogue system that O'Neill's system could be used for presenting media to a user. (EX-1003, ¶¶353-56.)

S. Claim 21

1. Determining a Reservation For the User in Response to One or More Requests [21A]

The O'Neill-Jones combination discloses this limitation.

As discussed above in §IV.C.2, O'Neill discloses a system that assists users with making various types of reservations, such as hotels, in response to user requests. (EX-1005, [0274]-[0342]; EX-1003, ¶¶358-359.)

Jones also discloses assisting users with locating restaurants, which may include determining a reservation to be made on behalf of the user. As discussed above, when a user requests information pertaining to "restaurants in Indiana" and the system uses its database of advertisements and the user's GPS coordinates to select from restaurant advertisements in Indiana to provide one in or near the designated geographic location. (EX-1006, [0031]; EX-1003, ¶¶357-60.)

2. Determining Promotional Content Based on Information Related to the Reservation [21B]

The O'Neill-Jones combination discloses this limitation.

As discussed above in §IV.C, Jones discloses determining targeted promotional content based on user requests and activities, such as assisting users in locating restaurants. (EX-1006, [0030]; EX-1003, ¶362.) A POSITA would therefore combine Jones's targeted advertising related to restaurants with O'Neill's booking reservations teachings. (EX-1003, ¶363.) The O'Neill system analyzes information regarding the user's intent to make certain purchases related to activities, such as reserving a hotel room or purchasing a cinema ticket. (EX-1005, [0391]-[0403].) The activity of reserving a hotel provides the O'Neill-Jones combination an opportunity for the system to “*determin[e] the promotional content based on information relating to the reservation*” and present the user with promotional content based on information related to the reservation. (EX-1003, ¶¶361-363.)

T. Claim 22

1. Determining a First Reservation in Response to User Request [22A]

The O'Neill-Jones combination discloses this limitation. (EX-1003, ¶¶364-368.)

As discussed above in §IV.S.1, O'Neill discloses a system that “*determine[es] a first reservation to be made on behalf of the user in response to*” a request, such as hotel reservations or reservations for cinema or theatre tickets. (EX-1005, [0274]-[0342], [0391]-[0408]; EX-1003, ¶¶364-67.) Also as discussed

above in §IV.S.1, Jones similarly discloses assisting users with locating restaurants, which may include determining a reservation to be made on behalf of the user. (EX-1006, [0031]; EX-1003, ¶368.)

O'Neill also provides an example where the user asks for assistance booking a hotel room and the system provides multiple hotel reservation options, including a second reservation option that is related to the first reservation option. (EX-1005, [0291]-[0327]; EX-1003, ¶¶364-68.)

2. Determining Promotional Content Based on Alternative Reservation [22B]

The O'Neill-Jones combination discloses this limitation.

As discussed above, O'Neill discloses that the system can provide “*a second reservation that is an alternative to the first reservation,*” such as alternative hotel reservation options or alternative room types. (EX-1005, [0290]-[0360], [0319]-[0325]; EX-1003, ¶374.)

The O'Neill-Jones combination also discloses “*determining the promotional content*” based on the alternative reservation. O'Neill discloses collecting information about the user's intent to make certain purchases. (EX-1005, [0366]-[0389].) Those user activities provide an opportunity for the O'Neill-Jones combination to present promotional content related to those activities, including providing promotional content relating to a second reservation that is an alternative to the first reservation.

Similarly, Jones discloses providing alternative promotional content from the businesses the user is searching for. (EX-1006, [0030] (providing user with alternative advertisements while user searches for a particular restaurant).) A POSITA would understand that the Jones system could present users with promotional content pertaining to the alternative restaurant reservation. (EX-1003, ¶¶369-74.)

U. Claim 23

The elements in limitations [23A]-[23B] are substantively identical to those in limitations [7A/41A/47A/51A]-[7B/41B/47B/51B]. The elements in limitations [23C] are substantively identical to those in claim [12]. For the reasons discussed above in §IV.M, the O’Neill-Jones combination discloses these limitations. (EX-1003, ¶¶375-383.)

V. Claim 30

1. Obtaining User Profile Information [30A]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶384-392.)

Jones discloses “obtaining” “*user profile information associated with the user*” that includes “*prior user interactions with items*,” namely tracking user engagement with promotional content, including “whether the user ‘clicked through’ or accessed the advertisement, whether the user bought products or requested services from an advertiser’s website referred to in the advertisement.”

(EX-1006, [0032]-[0033]; EX-1003, ¶¶385.) Obtaining and using user profile information was well-known as evidenced by Howard (discussed below in Ground 4). Howard discloses a “profile data store,” “which may contain user profile information, such as data on what the user has previously requested” and is used “to guide the dialog manager and the associated filtering function in identifying information the user may or may not be interested in.” (EX-1009, 6:31-54.)

The O’Neill system similarly stores information regarding the user’s interactions with the system and the system’s interactions with the user to improve the dialogue with the user and allows the dialogue to evolve dynamically. (EX-1005, [0064], [0076]-[0077].) This discourse knowledge is used by the system to prepare responses to the user. (EX-1005, [0103]-[0104]; EX-1003, ¶¶388-389.)

2. Identifying Categories of Items Based on Prior User Interactions [30B]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶393-404.)

Jones discloses tracking user interactions (“*user profile information*”) with promotional content to provide targeted promotional content. For example, Jones discloses that the system “may maintain a record of advertisements presented in response to queries containing particular keyword(s) and whether users interacted with the advertisements.” (EX-1006, [0032]-[0033].) The system then accesses that information database when it selects an advertisement for the user. (EX-1006,

[0031].) For example, where a user requests information pertaining to “restaurants in Indiana” and the system presents restaurant advertisements in Indiana to the user based on prior engagement with the advertisement and the user’s geographic location. (EX-1006, [0031].) The system thus identifies “*categories of items based on the prior user interactions.*” (EX-1003, ¶¶395-398.)

Moreover, a POSITA would be motivated to modify O’Neill’s user tracking process (discussed above in §IV.V.1) with Jones to track user interactions with promotional content, including “*identifying*” “*categories of items based on prior user interactions specified by the user profile information.*” (EX-1003, ¶¶399-403.) The O’Neill system tracks all user interactions with the system so it would automatically track user interactions with promotional content and identify corresponding categories of items associated with those interactions. (EX-1003, ¶403.)

3. Determining a Promotional Item Associated with One or More Categories [30C]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶405-410.)

As discussed above in §IV.V.2, the O’Neill-Jones combination would “*identify[] categories of items based on prior user interactions specified by the user profile information,*” and use that information to determine and provide the user with promotional content associated with one or more of those categories.

Jones discloses a system that uses a number of factors to weight and determine the appropriate “*promotional item associated with one of the categories.*” (EX-1006, [0029].) In one example, the user tells the system he is interested in attending a performance at The Kennedy Center for the Performing Arts and in response the system determined promotional item—advertisements related to Broadway shows— associated with the category of theatre. (EX-1006, [0029]; EX-1003, ¶¶407-408.)

Similarly, Jones provides an example in which the user requests information related to the category of restaurants and the system determines a promotional item—advertisements for restaurants—associated with the category of restaurants. (EX-1006, [0031]; EX-1003, ¶409.)

W.Independent Claim 50

1. [50P]-[50E] and [50G]

The elements in limitations [50P]-[50C] and [50G] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitations [50D]-[50E] are substantively identical to those in limitations [30A]-[30B].

For the same reasons discussed above in §§IV.B.1-4; IV.B.12; and §§IV.V.1-2, the O’Neill-Jones combination discloses these limitations. (EX-1003, ¶¶411-428, 436-438.)

2. Determining Promotional Content Associated with One of the Categories Based on the First and Second Natural Language Utterances [50F]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶429-435.)

As discussed above in §IV.V.3, the O’Neill-Jones combination discloses “*determining a promotional item associated with one of the categories,*” which is substantially similar to this claim limitation which requires “*determining...promotional content associated with one of the categories,*” “*based on the first natural language utterance and the second natural language utterance.*” (See EX-1005, [0213]-[0214].) This claim limitation additionally requires determining the promotion content based on both the first and second natural language utterances.

As discussed above in §IV.C, O’Neill discloses using all of the user’s utterances and the system’s responses to process future user utterances. (EX-1005, [0103]-[0104]; EX-1003, ¶¶403-431.) Therefore, the O’Neill-Jones combination would use all information when “*determining ... promotional content associated with one of the categories,*” including “*the first natural language utterance and the second natural language utterance.*”

Jones discloses a system that uses a number of factors to weight and determine the appropriate “*promotional content,*” including prior “requests

submitted” by the user. (EX-1006, [0032]; *see also* EX-1006, [0029].) As Jones’s system is already designed to determine promotional content based on prior requests and the current utterance, the system could also be used to determine promotional content based on both “*the first natural language utterance and the second natural language utterance.*” (EX-1003, ¶¶432-435.)

X. Independent Claim 53

The elements in limitations [53P]-[53C] and [53G] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L]. The elements in limitations [53D]-[53F] are substantively identical to those in limitations [30A]-[30C]. Claim limitation [53F] is also substantively identical to limitation [50F]. For the same reasons discussed above in §§IV.B.1-4; IV.B.12; IV.V; IV.W.3, the O’Neill-Jones combination discloses all limitations in claim 53. (EX-1003, ¶¶439-462.)

Y. Independent Claim 54

1. [54P]-[54E] and [54H]-[54I]

The elements in limitations [54P]-[54C] and [54I] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [54D] are substantively identical to those in limitation [30A].

The elements in limitation [54E] are substantively identical to those in limitation [16B].

The elements in limitation [54H] are substantively identical to those in limitation [50F]. For the same reasons discussed above in §§IV.B.1-4; IV.B.12; IV.V.1; IV.N.2; IV.W.3; the O’Neill-Jones combination discloses all limitations in claim 54. (EX-1003, ¶¶463-491.)

2. Determining Applications for Processing Requests [54F]

The O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶481-482.)

O’Neill discloses a system that associates user requests with applications. In one example, a user first asks the system to assist in the booking of a hotel accommodation and the system processes that request using the domain specific “AccommodationExpert,” an application which assists the user with the hotel booking. (EX-1005, [123].) O’Neill further discloses that the demands of specific applications will determine how the system processes the user requests, as the system improved dialogue manager can be used to handle various requests. (EX-1005, [0123].) To the extent the AccommodationExpert is not an “application,” O’Neill further discloses that “[i]t will of course be appreciated that other service and support agents could be included into the improved dialogue manager 20 in accordance with the demands of specific applications.” (EX-1005, [0118].)

3. Identifying Categories of Items Related to Applications Based on Prior Use Interactions [54G]

The O'Neill-Jones combination discloses this limitation. (EX-1003, ¶¶483-485.)

As discussed above in §IV.W.3, O'Neill discloses identifying “*categories of items based on the prior user interactions*” as the system assists with booking cinema tickets or hotel rooms using user profile information. Those categories (cinema tickets and hotel accommodation) “*relate to the one or more applications.*” (EX-1005, [0206], [0294], [0394].)

Z. Independent Claim 55

The elements in limitations [55P]-[55C] and [55I] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [55D] are substantively identical to those in limitation [30A].

The elements in limitation [55E] are substantively identical to those in limitation [16B].

The elements in limitations [55F]-[55G] are substantively identical to those in limitations [54F]-[54G].

The elements in limitation [55H] are substantively identical to those in limitation [50F] and [30C].

For the same reasons discussed above in §§IV.B.1-4; IV.V.12; IV.Y2-3; IV.V.3; IV.N.2; IV.W.3, the O’Neill-Jones combination discloses each limitation in claim 55. (EX-1003, ¶¶492-522.)

V. GROUND 2: O’Neill-Jones-Franco Combination Renders Obvious Claims 13-15, 24, 33-38

A. Combination Overview

1. Franco

Franco describes, through multiple embodiments, methods and systems for “correcting errors in speech recognition systems.” (EX-1007, [0002].)

One embodiment involves a scenario where a user offers an utterance, the utterance is incorrectly recognized, and the user repeats the same utterance without prompting. (EX-1007, [0018].) In such scenario, the method combines hypothesized interpretations of the first and second (repeated) utterance to maximize the likelihood of correct interpretation. (EX-1007, [0020] (“By assuming that a subsequent speaker utterance will comprise all or part of a previously spoken **(and incorrectly recognized)** utterance, the acoustic evidence from the previous and subsequent utterances can be combined in an efficient manner to produce a more accurate corrected recognition.”), [0006]; EX-1003, ¶¶525-526.)

2. Motivation to Combine

A POSITA would have been motivated to integrate Franco's speech recognition error correction features with O'Neill's interactive dialogue system and Jones's targeted advertising system for a number of reasons. (EX-1003, ¶527.)

First, O'Neill recognizes that "various modifications and variation can be made" to the O'Neill system. (EX-1005, [0423].) A POSITA would understand O'Neill expressing a motivation to combine its system with systems like Franco, that could enhance the system in O'Neill to add correction of errors in speech recognition. O'Neill also recognizes that its system builds on the DARPA Communicator system architecture but does not provide all of the details of the operation of that system. (EX-1005, [0005]-[0019].) A POSITA would therefore be motivated look to other references, like Franco, to implement parts of O'Neill's system, like the speech recognizer and dialogue manager. (EX-1003, ¶¶528.)

Second, Franco also recognizes the motivation, stating that there is a "need in the art for a method and apparatus for error correction in speech recognition applications." (EX-1007, [0005].) A POSITA would understand that Franco is suggesting that its method of error correction was broadly applicable and could be implemented and combined with speech recognition applications, like O'Neill and Jones. (EX-1003, ¶529.)

The results of the combination would have been predictable and a POSITA would have had a reasonable expectation of success because Franco discloses that its error recognition module can be “implemented as a physical device or subsystem that is coupled to a processor through a communication channel.” (EX-1007, [0027].) O’Neill and Jones are implemented through the telephone system, which contains communication channels. (EX-1003, ¶¶530-531.)

B. Claim 13

1. [13A]-[13B]

The elements in limitation [13A] are substantively identical to those in limitation [2A].

The elements in limitation [13B] are substantively identical to those in claim [8].

For the same reasons discussed above in §§IV.C.1; IV.I, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶¶532-537.)

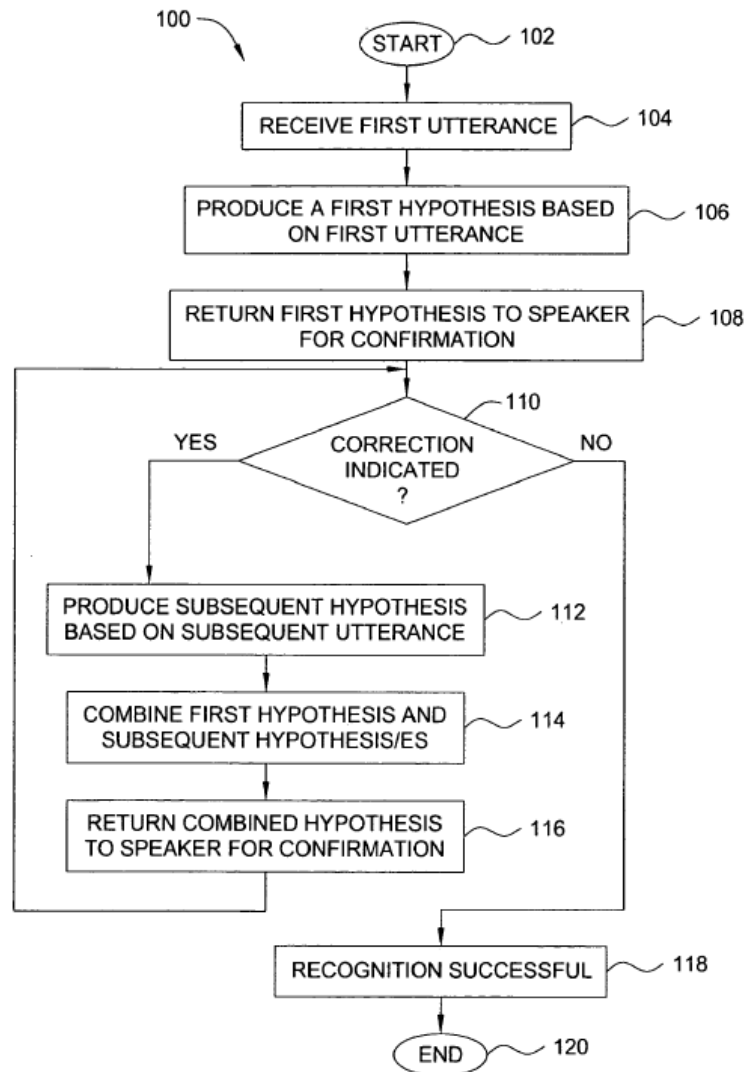
2. Processing the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [13C]

The O’Neill-Jones-Franco combination discloses this limitation.

Franco discloses a “method for recognizing user speech includes receiving a first utterance from the user, receiving a subsequent utterance from the user, and combining acoustic evidence from the first utterance with acoustic evidence from

the subsequent utterance in order to recognize the first utterance.” (EX-1007, [0006].)

In one embodiment involving successive utterances (shown in Figure 1, below), the system receives an utterance (104) and provides a hypothesized response to the speaker (108). (EX-1007, Figure 1.)

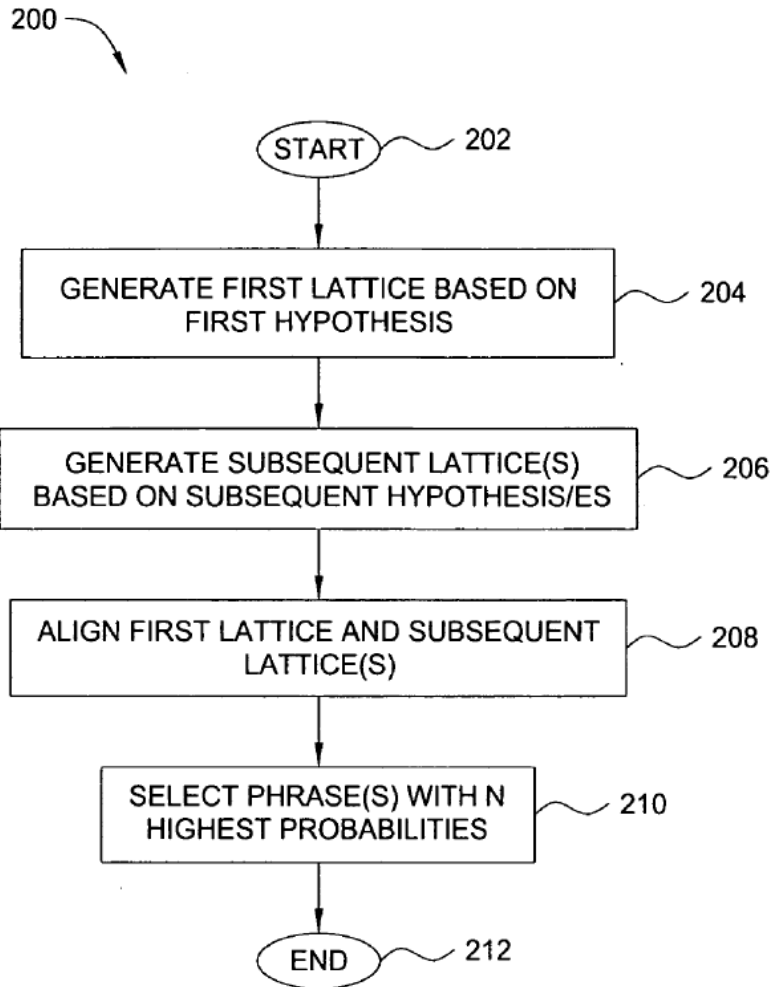


Franco, Figure 1

If the subsequent natural language utterance indicates a correction is needed, the system then generates a hypothesis for the second utterance **and a new, combined hypothesis for the first utterance** based on the first and second utterances. (EX-1007, [0017].)

In this embodiment where the speaker repeats all or part of the same utterance “without an intervening prompt from the method 100,” a hypothesis is presented “to the speaker for review/confirmation.” (EX-1007, [0018]-[0019].) If “further correction to the combined hypothesis is necessary,” the system will “continue to combine hypotheses in an effort to correct recognition **of the first utterance** until recognition is determined to be successful” as shown in Figure 1 above. (EX-1007, [0018]; EX-1003, ¶¶545-547.)

In another embodiment (Figure 2, below), Franco discloses combining hypotheses by first generating word lattices corresponding to the hypotheses.



Franco, Figure 2

After aligning lattices, the system can “select the phrase(s) with the N highest probabilities as the combined hypothesis to be presented to the speaker” (EX-1007, [0026].) Alternatively, the highest probability hypothesis can be presented to the speaker and if the speaker indicates that the hypothesis is incorrect, “then the method 200 selects the phrase with the second highest probability as an alternative combined hypothesis.” (EX-1007, [0026].)

In this example, the system reinterprets the misrecognized words from the first utterance in response to subsequent utterances. (EX-1007, [0020]; EX-1003, ¶¶550.) This example demonstrates that Franco discloses a system that “*determine[s] a reinterpretation*” of a prior natural language based on a “*subsequent natural language utterance.*” (EX-1003, ¶¶550.) Specifically, Franco discloses using a subsequent utterance (a “*third natural language utterance*”) to “*determine a reinterpretation of*” a prior natural language utterance (“*the second natural language utterance*”).

This disclosure is consistent with the ’536 patent intrinsic record which states that “conversational language processor 120 may support adaptive misrecognition to reinterpret a current utterance and/or one or more previous utterances.” (EX-1001, 5:4-6.) Rather than disclose structure for the “adaptive misrecognition” used to interpret utterances in the specification, the applicants referenced “U.S. patent application Ser. No. 11/200,164, ... which issued as U.S. Pat. No. 7,620,549” [“Di Cristo”]. (EX-1001, 5:14-25-64.) Di Cristo explains that one way the “adaptive misrecognition engine” works is by running through a list of highest scored interpretations and presenting them serially to the user. (EX-1012, 33:42-48 (when the “adaptive misrecognition engine determines that the interpretation generated” is incorrect, the system should “select the next best interpretation”).) Franco thus discloses a process for reinterpreting misrecognized

utterances that mirrors the support for the alleged invention in the '536 patent.

(EX-1003, ¶¶553-554.)

O'Neill similarly discloses a system that includes what O'Neill calls the DiscourseManager, which uses confirmation strategy and dialogue to determine if the user's current utterance repeats, modified or negates prior information provided by the user, including repairing information provided by the user. (EX-1005, [0212]-[0213]; *see also* EX-1005, [0236].) O'Neill's disclosure of using subsequent natural language utterances, including a third natural language utterance, to determine if the prior natural language utterances were interpreted correctly by the system and then determines whether that prior information (or interpretation) needs to be repaired is "*determin[ing] a reinterpretation of the second natural language utterance*" based on the "*third natural language utterance*." (EX-1003, ¶¶538-554.)

C. Claims 14 and 34

The O'Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶¶555-561, 629-631.)

The elements in claims [14] and [34] are substantively identical to those in limitations [1K]/[32K]. For the same reasons discussed above in §IV.B.11, the O'Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶556.)

The only difference between claims [14] and [34] and [1K]/[32K] is that claims [14] and [34] require determining the promotional content based on the “reinterpretation” rather than the “interpretation” of the second natural language utterance. However, this difference is immaterial, as the O’Neill-Jones-Franco combination would “*determin[e] the promotional content*” based on the interpretation of the utterance that is operative at the time the promotional content is generated. (EX-1003, ¶¶555-561.) Therefore, once the O’Neill-Jones-Franco combination has reinterpreted the “second natural language utterance,” the system would necessarily use that reinterpretation to determine promotional content for the user. (EX-1003, ¶¶555-561.) As discussed above in §IV.B.11, O’Neill discloses using all information regarding the current conversation to determine responses to users, including determining promotional content in the combined system.

D. Claims 15 and 35

1. [15A]/[35A]

The elements in limitations [15A]/[35A] are substantively identical to those in limitation [1K]. For the same reasons discussed above in §IV.B.11, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶¶562-564.)

2. Determining Other Promotional Content Based on the Reinterpretation [15B]/[35B]

The O’Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶¶565-569.)

The elements in claim limitations [15B]/[35B] are substantively identical to those in limitations [1K]/[32K]. For the same reasons discussed above in §IV.B.11, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶566.)

The only difference between claim limitations [15B]/[35B] and [1K]/[32K] is that claim limitations [15B]/[35B] require determining the “other” promotional content based on the “reinterpretation” rather than determining promotional content based on the “interpretation” of the second natural language utterance. However, this difference is immaterial. As discussed above, the O’Neill-Jones-Franco combination would “*determin[e] the promotional content*” based on the interpretation of the utterance that is operative at the time the promotional content is generated. (EX-1003, ¶¶567-659.) Therefore, if the O’Neill-Jones-Franco combination determined promotional content has reinterpreted the “second natural language utterance,” the system would necessarily use that reinterpretation to determine promotional content for the user. (EX-1003, ¶¶565-569.) As discussed above in §IV.C.11, O’Neill discloses using all information regarding the current

conversation to determine responses to users, including determining promotional content in the combined system.

3. Presenting the Promotional Content to the User [15C]/[35C]

The O’Neill-Jones-Franco combination discloses this limitation.

The elements in limitations [15C]/[35C] are substantively identical to those in limitations [1L]/[32L]. For the same reasons discussed above in §IV.B.12, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶571.)

The only difference between claim limitations [15C]/[35C] and [1L]/[32L] is that claim limitations [15C]/[35C] require “*presenting*” “*the other promotional content*” that was determined in claim limitations [15B]/[35B] “*to the user.*” However, this difference is immaterial, as the O’Neill-Jones combination would present any promotional content the system determines is appropriate for the user. (EX-1003, ¶¶570-571.)

E. Claim 24

1. Determining at Least One Request is Incomplete or Ambiguous [24A]

The O’Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶¶572-582.)

Each of O’Neill, Jones and Franco disclose a system that can determine if “*one or more requests is incomplete or ambiguous.*” As discussed above in §V.B.2, O’Neill’s DiscourseManager uses confirmation strategy and dialogue to

determine if the user's current utterance repeats, modified or negates prior information. (EX-1005, [0212]-[0213]; *see also* EX-1005, [0236].) In other words, if the user's utterance repeats, modified or negates prior information, then the DiscourseManager determines that the prior utterance was incomplete or ambiguous and then uses its confirmation strategy to confirm its interpretation. (EX-1003, ¶¶573-574.)

Jones also discloses a system that can determine if a request is incomplete or ambiguous. Jones recognizes that some words, like "apple" may have multiple meaning or connotations to a user and qualification may be necessary. (EX-1006, [0058].) Therefore, Jones teaches that if a term is ambiguous, then the system can prompt the user to clarify the meaning or the system can attempt to determine the meaning itself. (EX-1006, [0053].)

As discussed above in §V.B.2, Franco teaches a system that is capable of determining that a user's utterance was incorrect, including ambiguous or incomplete. (§V.B.2.) Franco discloses a method of then correcting the user's utterance using information the system obtains in subsequent user utterances. (§V.B.2.)

2. Monitoring Interactions of the User with the Promotional Content [24B]

The O'Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶¶583-590.)

Jones's targeted advertising system includes a robust process for
“*monitoring... interaction of the user with the promotional content,*” including “the number of times an advertisement has been presented, the length of time a user (or requester) viewed an advertisement, whether the user ‘clicked through’ or accessed the advertisement, whether the user bought products or requested services from an advertiser’s website referred to in the advertisement” (EX-1006, [0032]-[0033].)

As discussed above in §V.B.2, O’Neill’s system also monitors the user’s interactions with the system and the system’s interactions with the user to improve the dialogue with the user. O’Neill discloses that the dialogue manager continues to use information from prior parts of the conversation with the user. (*See, e.g.,* EX-1005, [0062], [0076]-[0077], [0099]-[0100].)

3. Interpreting the Incomplete or Ambiguous Request Based on the Interaction [24C]

The O’Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶¶591-596.)

While O’Neill does not explicitly disclose interpreting an incomplete or ambiguous request based on an interaction of a user with promotional content, the O’Neill system does use all information from the interactive dialogue session with the user to determine the interpretation and response to the user’s utterances, including user interactions with the promotional content when combined with

Jones. As discussed above in §IV.B, O'Neill contains “components for storing information acquired or generated during a dialogue session,” which in a combined O'Neill-Jones system would include generated promotional content and user interactions with that promotional content. (See EX-1005, [0103].) This data forms the system's “discourse knowledge,” which is used by the system to “prepare system utterances in accordance with their generic and domain specific rules.” (EX-1005, [0103]-[0104]; EX-1003, ¶¶592-596.) The system can then “retrieve information collected at an earlier point in a dialogue to assist in a current query,” meaning that prior information generated by the system or user during a dialogue session—like user interaction with promotional content—is used to assist in interpretation and processing of future user requests, such as the “*interpreting ... the at least one incomplete or ambiguous request.*” (See EX-1005, [0076]-[0077]; EX-1003, ¶¶592-596.)

As discussed above in §V.B.2, Franco discloses a “method for recognizing user speech includes receiving a first utterance from the user, receiving a subsequent utterance from the user, and combining acoustic evidence from the first utterance with acoustic evidence from the subsequent utterance in order to recognize the first utterance.” (EX-1007, [0006].) The “subsequent utterance” used by Franco to interpret the prior utterance is not limited to any particular type of utterance. (EX-1003, ¶¶542-551.) Therefore, the “subsequent utterance” used by

Franco to interpret the prior utterance can include an “*interaction of the user with the promotional content*.” Similarly, as discussed above in §V.E.1, the prior utterance can include an “*incomplete or ambiguous request*.”

Similarly, Jones uses context information to interpret words with multiple meanings like “apple” in the example above discussed in §IV.D.1. (EX-1003, ¶¶576-579.)

F. Independent Claim 33

The elements in limitations [33P], [33B], [33D], [33E] and [33I] are substantively identical to those in limitations [1P]/[32P], [1B]/[32B], [1D]/[32D], [1F]/[32F], [1L]/[32L].

The elements in limitation [33A] are substantively identical to those in limitation [2A].

The elements in limitation [33C] are substantively identical to those in limitations [1C] and [2A].

The elements in limitations [33F] and [33G] are substantively identical to those in limitations [13B] and [13C].

The elements in limitation [33H] are substantively identical to those in limitation [1K] (interpretation) and claim [14] (reinterpretation).

For the same reasons discussed above in §§IV.B; IV.C; V.B; V.C, the O’Neill-Jones-Franco combination discloses all elements in claim 33. (EX-1003, ¶¶597-628.)

G. Independent Claim 36

1. [36P]-[36D], [36F]-[36I]

The elements in limitations [36P]-[36D] and [36I] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitations [36F]-[36G] are substantively identical to those in limitations [13B]-[13C].

The elements in limitation [36H] are substantively identical to those in limitation [33H].

For the same reasons discussed above in §§VI.B; V.B; V.F, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶¶641-655, 659-670.)

2. [36E]

The elements in limitation [36E] are substantively identical to those in limitation [1F]. For the same reasons discussed above in §IV.B.7, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶¶657-658.)

The only difference between claim limitation [1F] and [36E] is that claim limitation [1F] requires “*processing ... based on the domain information*” while claim limitation [36E] requires the system to “*process ... based on the first natural*

language utterance.” However, as discussed above in in §IV.B.7, the “domain information” is determined based on the first natural language utterance.

H. Independent Claim 37

1. [37P]-[37E] and [37H]-[37I]

The elements in limitations [37P]-[37C] and [37G] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [37D] are substantively identical to those in limitation [16B].

The elements in limitations [37E], [37H] and [37I] are substantively identical to those in limitations [24A], [24B] and [24C].

For the same reasons discussed above in §IV.B, the O’Neill-Jones-Franco combination discloses these limitations. (EX-1003, ¶¶671-688, 692-700.)

2. Determining Promotional Content Relating to Request [37F]

The elements in limitation [37F] are substantively identical to those in limitation [1K]. For the same reasons discussed above in §IV.B.11, the O’Neill-Jones-Franco combination discloses this limitation. (EX-1003, ¶690.)

The only difference between claim limitation [37F] and [1K] is that claim limitation [37F] requires the system to determine “*the promotional content that relates to the one or more requests,*” associated with the second natural language utterance whereas claim limitation [1K] requires the system to determine “*promotional content based on the interpretation*” of the second natural language

utterance. This difference is immaterial as the second natural language utterance used by the O’Neill-Jones combination to determine promotional content in claim 1 also can contain “*one or more requests*” that the system would use to determine the promotional content. (EX-1003, ¶¶690-691.)

I. Independent Claim 38

The elements in limitations [38P]-[37C] and [37G] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [38D] are substantively identical to those in limitation [16A].

The elements in limitation [38E] are substantively identical to those in limitation [24A].

The elements in limitation [38F] are substantively identical to those in limitations [1K] and [37F].

The elements in limitations [38H] and [38I] are substantively identical to those in limitations [24B] and [24C].

For the same reasons discussed above in §§IV.B; IV.N; V.E, the O’Neill-Jones-Franco combination discloses all limitations in claim 38. (EX-1003, ¶¶701-730.)

**VI. GROUND 3: O’Neill-Jones-Wang Combination Renders Obvious
Claims 25-26, 31, 39-43**

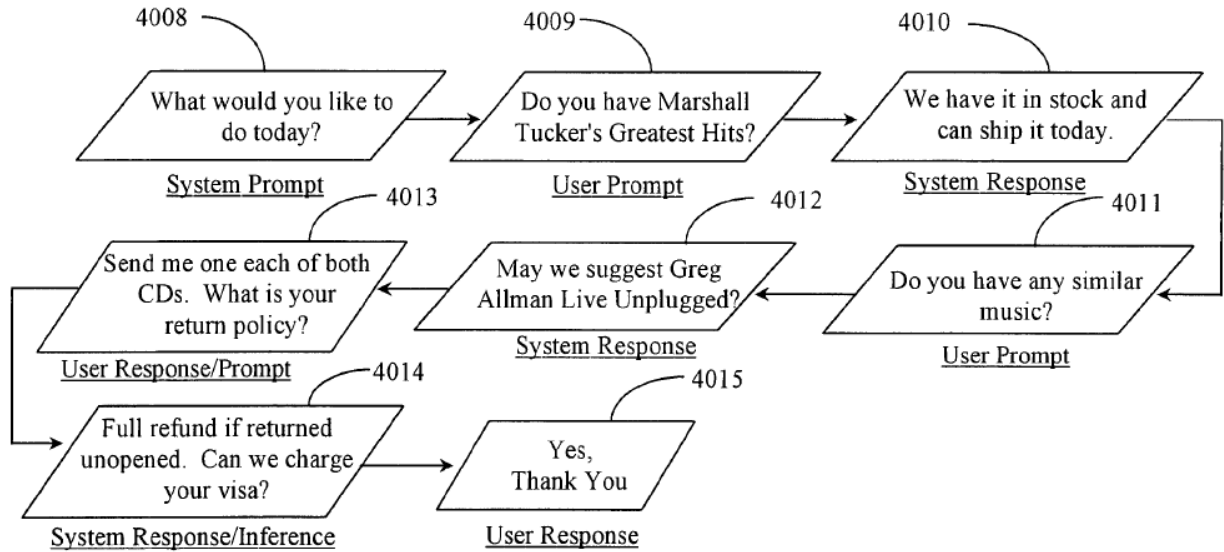
A. Combination Overview

1. Wang

Wang describes a “voice application creation and deployment system” that includes a “voice application server for creating and serving voice applications.” (EX-1008, Abstract.) In exemplary embodiments, voice applications are deployed in phone-based customer service situations such as directory assistance, banking and purchasing goods. (*See* EX-1008, 53:34-56:42.)

“By providing multi-slot capability and a rich semantic base for voice recognition much more work can be accomplished using less request/response interaction.” (EX-1008, 55:50-52.) This contrasts with the prior art where systems were unable to process two requests in a single utterance. (EX-1008, 55:29-48.)

Figure 40B (below) depicts a “system/client interaction dialog” using the Wang system in trying to purchase music CDs over the phone. (EX-1008, 55:53-55.) “At step 4012, the system makes a suggestion to the client to the effect of ‘May we suggest Greg Allman Live Unplugged?’” (EX-1008, 56:12-14.) “At step 4013 the client responds by asking the company to send one each of both CDs, and asks about the return policy in the same dialog.” (EX-1008, 56:16-18.) The system is thereafter able to process the transaction and answer the question regarding return policy in parallel. (EX-1008, 56:18-23.)



Wang, Figure 40B

2. Motivation to Combine

A POSITA would have been motivated to combine Wang's functionality for processing multi-request utterances into the O'Neill-Jones combination for a number of reasons. (EX-1003, ¶376.)

First, O'Neill recognizes that it is "apparent to those skilled in the art that various modifications and variation can be made" to the O'Neill system. A POSITA would understand O'Neill expressing a motivation to combine its system with systems like Wang, that could enhance the system in O'Neill to understand multi-request utterances in speech recognition. (EX-1003, ¶737.) Second, Wang likewise motivates the combination. Wang discloses the benefit of being able to process multi-request utterances in customer service applications, stating, in an ordering context, that the lack of such capability "requires 3 request/response

interactions before an order can be successfully created.” (EX-1008, 55:37-49.) A POSITA would therefore be motivated to incorporate Wang’s teachings into the O’Neill-Jones combination which includes phone-based directory assistance (Jones) and a dialog manager (O’Neill). (EX-1003, ¶738.)

The results of the combination would have been predictable to a POSITA and the POSITA would have had a reasonable expectation of success because Wang discloses that its features can be broadly implemented in “serving voice applications to clients over a communication network.” (EX-1008, 6:38-45; EX-1003, ¶739.)

B. Claim 25

1. First Request Associated with First Application and Second Request Associated with Second Application [25A]

The O’Neill-Jones-Wang combination discloses this limitation. (EX-1003, ¶¶740-744.)

Wang discloses a single utterance containing a “*first request*” and a “*second request*.” As noted in the example above, the user asks the seller “to send one each of both CDs, and asks about the return policy **in the same dialog**.” (EX-1008, 56:16-18.)

A POSITA would understand that the first and second requests are “*associated with a first application*” and a “*second application different than the first application*,” respectively.

A user calls a server “for interaction.” (EX-1008, 50:50.) An “[a]pplication server” is a “hosted server” to facilitate “creation, configuration and deployment of voice applications designed for ... completing real time transactions representing business between a client and a product/service provider.” (EX-1008, 51:4-9.) The system “may deploy or distribute one or more application servers as may be desired.” (EX-1008, 52:35-39.)

In Wang, after received the exemplary utterance above, “order preparation can take place in the background while the system addresses” the question about the return policy. (EX-1008, 56:21-23.) A POSITA would understand that such functionality would be implemented by accessing two different applications—one relating to order processing and one that includes basic return policies. (EX-1003, ¶742.) Such architecture was well known as evidenced by U.S. Patent 5,321,608 (“Namba”). Namba discloses natural language processing of a single utterance containing multiple requests such as “Show sales in a list, and represent its contents by a bar graph.” (EX-1013, 23:17-18.) Using the disclosed system a user “can adequately operate specific applications through the natural language without knowledge of complicated operation instructions, language system, operating instructions system, and data structure system **which are different for each application.**” (EX-1013, 23:46-51; EX-1003, ¶¶744.)

As discussed above in §IV.Y.1, O’Neill also discloses a system that associates user requests with applications. (*See also* EX-1003, ¶¶740-741.)

2. Promotional Content Related to One or More of the First or Second Request [25B]

The O’Neill-Jones-Wang combination discloses this limitation for the same reasons set forth in limitations [1K]/[32K]/[1L]/[32L]/[16B] which note that the O’Neill-Jones combination discloses determining and presenting promotional content associate with a request. (EX-1003, ¶¶745-456.)

C. Claims 26, 40

The O’Neill-Jones-Wang combination discloses this limitation.

Wang discloses executing a multi-request utterance that includes one request associated with “*a commerce application*” (e.g., purchasing a CD) and another request associated with a database application (providing information about return policies) which is an “*application type different than the first application*”. (EX-1003, ¶¶748-749.)

D. Claim 31

1. [31A]-[31B] and [31E]

The elements in limitations [31A] and [31E] are substantively identical to those in limitations [30A] and [30C].

The elements in limitation [31B] are substantively identical to those in limitation [16B].

For the same reasons discussed above in §§IV.V; IV.N, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶750-756, 767-769.)

2. Determining Application for Processing One or More Requests [31C]

The O’Neill-Jones-Wang combination discloses this limitation.

As discussed above in §VI.B.1, Wang and O’Neill associate user requests with applications. The only difference in claim limitation [31C] is that it requires determining the application for processing the request. However, in order for the system to “associate” the request with an application in claim limitation [25A] it would necessarily have “determine[ed] ... one or more applications for processing the one or more requests.” For the same reasons discussed above in §VI.B.1, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶757-761.)

3. Identifying Categories of Items Based on Prior Use Interactions Related to One or More Applications [31D]

The elements in limitation [31D] are substantively identical to those in limitation [30B]. For the same reasons discussed above in §IV.V, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶762-764.)

The only element of this claim limitation that does not appear in claim limitation 30B is “*wherein the categories relate to the one or more applications.*” To the extent this element is not already contained within the discussion above

regarding claim limitation [30B], this element is independently disclosed in Jones.

In one of the examples, Jones discloses the system identifying the category of electronic merchandise, such as MP3 songs, based on prior user interaction. (EX-1006, [0055].) A POSITA would understand that the category of MP3 songs relate to one or more applications, such as a music player. (EX-1003, ¶765.)

As discussed above in §IV.V, O’Neill discloses identifying “categories of items based on the prior user interactions” as the system assists with booking cinema tickets or a hotel room by obtaining user profile information about the user. Those categories (cinema tickets and hotel accommodation) relate to one or more applications. (EX-1005, [0206], [0294], [0394].) While O’Neill does not disclose using this user profile information to determine promotional content, the O’Neill-Jones combination would determine promotional content and promotional items based on the categories disclosed in O’Neill that relate to one or more applications. (EX-1003, ¶¶762-766.)

E. Independent Claim 39

The elements in limitations [39P]-[39C] and [39F] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitations [39D]-[39E] are substantively identical to those in limitations [25A]-[25B].

For the same reasons discussed above in §§IV.B; VI.B, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶770-787.)

F. Dependent Claim 41

The elements in limitations [41A]-[41B] are substantively identical to those in limitations [7AP]-[7B]. For the same reasons discussed above in §IV.H, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶788-790.)

G. Dependent Claim 42

The elements in claim [42] are substantively identical to those in limitations [12]. For the same reasons discussed above in §IV.M, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶791-793.)

H. Independent Claim 43

The elements in limitations [43P]-[43C] and [43F] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitations [43D]-[43E] are substantively identical to those in limitations [39D]-[39E].

For the same reasons discussed above in §§IV.B; VI.E, the O’Neill-Jones-Wang combination discloses these limitations. (EX-1003, ¶¶794-811.)

**VII. GROUND 4: O’Neill-Jones-Howard Combination Renders Obvious
Claims 27-29, 44-49**

A. Combination Overview

1. Howard

Howard describes a system that “allows the user to communicate in a natural way with a variety of devices communicating with the home network or home gateway” (EX-1009, Abstract.) using “speech recognition and natural language parsing” to provide “a fast and far less intrusive interface for a user in selecting a particular activity” (EX-1009, 1:66-2:10).

Howard uses a “natural language dialog module” to “access[] the electronic activity guide or guides, interpret[] the meaning of the user’s input speech and provide[] appropriate information to the network to retrieve specific information from any of the devices connected to the network.” (EX-1009, 2:24-30.) Howard provides a summary of the system architecture in Figure 1:

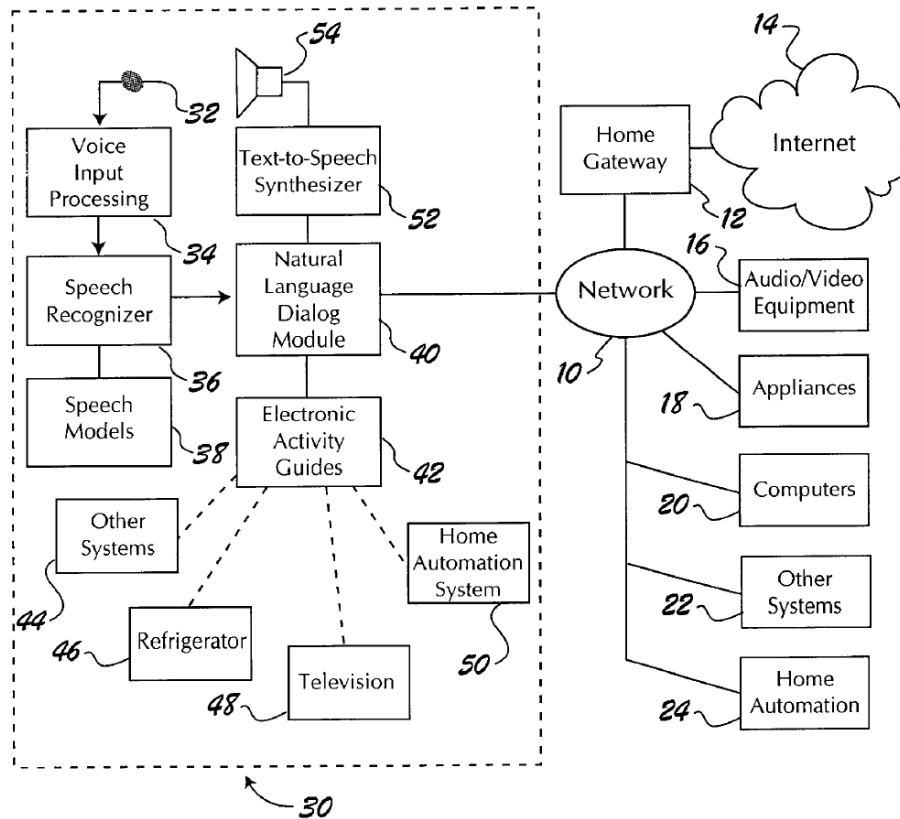


FIG. 1

Howard, Figure 1

User utterances are input from a microphone 32 and the supplied to voice input processing, which digitizes the speech. (EX-1009, 4:29-32.) The digitized speech is then input to the speech recognizer “to convert the spoken utterances into text form or symbolic form.” (EX-1009, 4:40-41.) “The output of speech recognizer 36 is fed to the natural language dialog module 40.” (EX-1009, 4:49-50.) The natural language dialog module is connected to the network 10, which connects the system to the devices 16-24. (EX-1009, 3:57-60.)

Each of the devices “is configured to work independent of its connection to the network.” (EX-1009, 3:57-64.) Exemplary devices include “audio/video equipment 16, appliances 18, computers 20 and other systems 22.” (EX-1009, 3:37-41.)

2. Motivation to Combine

A POSITA would have been motivated to combine O’Neill’s teachings regarding its interactive natural language dialogue system with Jones’ teachings of targeted promotional content and Howard’s teaching of a natural language dialog system that can be used to control multiple devices. (EX-1003, ¶818.)

A POSITA would have been motivated to make this combination for numerous reasons. First, as discussed above, a POSITA would have been motivated to make this combination because both O’Neill and Jones suggest that their systems can be combined with other speech processing systems. (*Supra* §IV.A.3.) Howard discloses a similar speech processing system as disclosed in O’Neill, as both references use domain-specific knowledge to interpret a user utterance. (EX-1009, 2:51-54 (“each appliance or other device communicating over the network has an associated electronic activity guide that conforms the dialog grammar to its particular domain.”).) (EX-1003, ¶¶819-820.)

Second, the systems disclosed in O’Neill, Jones and Howard are designed to assist users with requests using ASR. (EX-1006, [0023] (system designed to

“receive[] voice requests or queries submitted by users ... and return[] responses resulting from the processing.”); EX-1005, Abstract (system designed to “receive[] a communication from a user” process the communication “and provide[] a response to the user.”); EX-1009, Abstract (“system allows the user to communicate in a natural way with a variety of devices” and synthesizes “responses to user’s request.”).) As O’Neill, Jones and Howard are attempting to accomplish the same goal of receiving requests, processing requests and responding to requests, a POSITA would be motivated to combine their teachings. (EX-1003, 821.)

Third, adding the targeted advertising of Jones to the teachings of Howard would be intuitive to a POSITA. The Howard system collects information regarding categories like user’s grocery store shopping lists and dinner plans. (EX-1009, 7:47-54.) A POSITA would be motivated to use that valuable consumer purchasing information with the targeted advertising teachings of Jones to select more accurate promotional content to provide to users. (EX-1003, 822.)

Finally, the combination is nothing more than the application of a known technique (using Howard’s natural language system to control multiple devices) to a similar known system (O’Neill’s interactive dialogue system combined with Jones’ targeted advertising system) that operates in the same way (analyzing natural language utterances to interact with user’s requests). (EX-1003, ¶824.)

A POSITA would have had a reasonable expectation of success in the combination and the results of the combination would have been predictable. Combining O’Neill with Jones and Howard would merely integrate one component from Jones (targeted advertising) and one component from Howard (using the dialogue system to interact with multiple devices) into O’Neill’s interactive dialogue system. O’Neill, Jones and Howard are also in the same field of the ’536 patent—using speech recognition to facilitate user interactions. (See EX-1001, 2:22-25; EX-1005, Abstract, [0023]-[0027]; EX-1006, [0009]; EX-1009, 1:66-2:10; EX-1003, ¶¶818-825.)

B. Claim 27

1. First Request Processed by a First Device and Second Request Processed by Second Device [27A]

The O’Neill-Jones-Howard combination discloses this limitation. (EX-1003, ¶¶826-831.)

Howard provides an example of a single user utterance which includes multiple requests that are processed by a “*be processed by a first device associated with the user and ... a second device associated with the user.*” In the example, the user asks “What can I make for dinner?” (EX-1009, 2:37-39.) The system would then check the contents of the refrigerator and make suggestions about possible menus and what groceries may need to be acquired. In response the user says “I think I will make chicken stew tonight. Please print me a shopping list.” (EX-1009,

2:46-47.) That utterance contains two separate requests sent to two separate devices: (1) the user is requesting that the refrigerator (“*first device associated with the user*”) ascertain which items for chicken stew are not in the refrigerator (“*first request*”); and (2) the user is requesting that the printer (“*second device associated with the user*”) print the shopping list (“*second request*”). (EX-1009, 2:37-50.)

2. Promotional Content Relates to One or More Requests [27B]

The elements in limitation [27B] are substantively identical to those in limitation [25B]. For the same reasons discussed above in §VI.B.2, the O’Neill-Jones combination discloses this limitation. (EX-1003, ¶¶832-834.)

C. Claims 28, 45

The O’Neill-Jones-Howard combination discloses this limitation. (EX-1003, ¶¶835-837.)

Howard discloses a number of different devices that may be used in the system, including “audio/video equipment,” which a POSITA would understand to include a “*media player device*.” (EX-1009, 3:37-38.) Howard further discloses other “*device type[s] different than the first device type*,” such as “appliances 18, computers 20 and other systems 22.” (EX-1009, 3:37-38.)

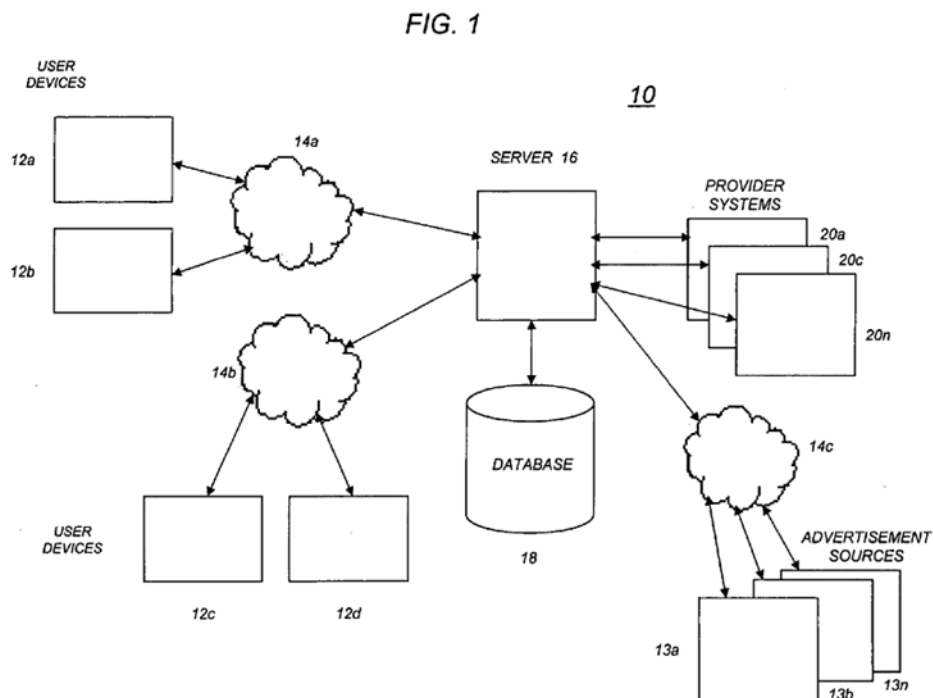
Also, Jones discloses that the system can use multiple device types that are different types, such as “mobile phones,” “voice-enabled PDA’s,” “radios,” “portable media devices,” etc. (EX-1006, [0023].)

D. Claims 29, 46

The O'Neill-Jones-Howard combination discloses this limitation. (EX-1003, ¶¶838-839.)

As discussed above, Howard discloses that each of the devices “has **independent processing capability** and at least some limited data storage ability” and “each device is **configured to work independent** of its connection to the network.” (EX-1009, 3:57-64.)

Similarly, Jones discloses in Figure 1 that multiple user devices 12a, 12b, 12c and 12d can connect independently to the servers 14a and 14b:



Jones, Figure 1

E. Independent Claim 44

The elements in limitations [44P]-[44C] and [44F] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [44D] are substantively identical to those in limitations [16A] and [27A].

The elements in limitation [44E] are substantively identical to those in limitation [39E].

For the same reasons discussed above in §§ VI.B; VI.N; VII.B; VI.E, the O’Neill-Jones-Howard combination discloses all limitations in claim 44. (EX-1003, ¶¶841-861.)

F. Dependent Claim 47

The elements in limitations [47A]-[47B] are substantively identical to those in limitations [7A]-[7B]. For the same reasons discussed above in § VI.H, the O’Neill-Jones combination discloses these limitations. (EX-1003, ¶¶862-867.)

G. Dependent Claim 48

The elements in claim [48] are substantively identical to those in limitations [12]. For the same reasons discussed above in § IV.M, the O’Neill-Jones combination discloses these limitations. (EX-1003, ¶¶868-870.)

H. Independent Claim 49

The elements in limitations [49P]-[49C] and [49F] are substantively identical to those in limitations [1P]-[1C] and [1L]/[32P]-[32C] and [32L].

The elements in limitation [49D] are substantively identical to those in limitation [44D].

The elements in limitation [49E] are substantively identical to those in limitation [39E].

For the same reasons discussed above in §§VI.B; VII.E; VI.E, the O’Neill-Jones-Howard combination discloses all limitations in claim 49. (EX-1003, ¶¶871-891.)

VIII. Mandatory Notices

A. Real Party In Interest

The real parties-in-interest are Samsung Electronics Co. Ltd. and Samsung Electronics America, Inc.

B. Related Matters

To the best of Samsung’s knowledge, the following is a list of judicial or administrative matter that would affect, or be affected by, a decision in the proceeding:

- *VB Assets, LLC v. Samsung Electronics Co. Ltd. et al.*, No. 2:24-cv-00828-JRG-RSP (E.D. Tex., filed Oct. 9, 2024)
- *VB Assets, LLC v. Amazon.com Servs. LLC*, No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019 2024)
- *VB Assets, LLC v. SoundHound AI, Inc.*, No. 1:24-cv-01279-MN (D. Del., filed Nov. 21, 2024)
- *VB Assets, LLC v. Apple, Inc.*, No. 1:24-cv-01368-MN (D. Del., filed Dec. 13, 2024)

- *Samsung Electronics Co. Ltd. et al. v. VB Assets, LLC*, IPR2025-00867 (PTAB)

C. Notice of Counsel and Service Information

Pursuant to 37 C.F.R. §§ 42.8(b)(3), 42.8, Petitioners designate the following lead and backup counsel:

Lead Counsel	Back Up Counsel
Lori Gordon (Reg. No. 50,633) GOODWIN PROCTER LLP 1900 N St. N.W. Washington, D.C. 20036 Phone: (202) 346-4000 Fax: (202) 346-4444 Gordon-ptab@goodwinlaw.com	Doug Kline (Reg No. 35,574) GOODWIN PROCTER LLP 100 Northern Avenue Boston, MA 02210 Tel: (617) 570 1000 Fax: (617) 523-1231 dkline@goodwinlaw.com
	Srikanth Reddy (<i>pro hac vice</i> application to be filed) GOODWIN PROCTER LLP 100 Northern Avenue Boston, MA 02210 Tel: (617) 570 1000 Fax: (617) 523-1231 sreddy@goodwinlaw.com
	Brian T. Drummond (Reg. No. 68,414) GOODWIN PROCTER LLP 100 Northern Avenue Boston, MA 02210 Tel: (617) 570 1000 Fax: (617) 523-1231 bdrummond@goodwinlaw.com
	Theodoros Konstantakopoulos (Reg. No. 74,155) GOODWIN PROCTER LLP 620 Eighth Avenue New York, NY 10018 Phone: (212) 813-8800 Fax: (212) 355-3333 tkonstantakopoulos@goodwinlaw.com

Petitioners consent to electronic service at the email addresses listed in the table above.

IX. Conclusion

IPR is respectfully requested.

Date: April 21, 2025

Respectfully submitted,

/Lori A. Gordon/

Lori A. Gordon (Reg. No. 50,633)
Goodwin Procter LLP
1900 N Street, N.W.
Washington, D.C. 20036
Phone: (202) 346 4435
Gordon-ptab@goodwinlaw.com

Lead Counsel for Petitioner

CERTIFICATE OF WORD COUNT UNDER 37 CFR §42.24(d)

Pursuant to 37 C.F.R. §42.24(a), Samsung hereby certifies that portions of the above-captioned Petition for *Inter Partes* Review of U.S. Patent No. 8,886,536 in accordance with and reliance on the word count provided by the word-processing system used to prepare this Petition, that the number of words in this paper is 13,982. Pursuant to 37 C.F.R. §42.24(a), this word count is in compliance and excludes the table of contents, table of authorities, mandatory notices under §42.8, certificate of service, certificate of word count, appendix of exhibits, and any claim listing. This word count was prepared using Microsoft Word.

Date: April 21, 2025

Respectfully submitted,

/Lori A. Gordon/

Lori A. Gordon (Reg. No. 50,633)
Goodwin Procter LLP
1900 N Street, N.W.
Washington, D.C. 20036
Phone: (202) 346 4435
Gordon-ptab@goodwinlaw.com

Lead Counsel for Petitioner

CERTIFICATE OF SERVICE

The undersigned certifies that a true copy of the Petition for *Inter Partes* Review of U.S. Patent No. 8,886,536 together with all exhibits identified in the above Table of Exhibits and Petitioners' Power of Attorney, have been served on the Patent Owner via FedEx Next Business Day Delivery on the below date, at the correspondence address of record as listed on the Patent Center:

Sheppard Mullin Richter & Hampton LLP
650 Town Center Drive
10th Floor
Costa Mesa, CA
Attorneys for VB Assets, LLC

Courtesy copies of the foregoing document were served by e-mail on the following counsel of record for Patent Owner in the concurrent litigation:

Theodore Stevenson, III (TX 19196650)
Brady Randall Cox (TX 24074084)
ALSTON & BIRD LLP
2200 Ross Avenue, Suite 2300
Dallas, TX 75201
Telephone: (214) 922-3400
Facsimile: (214) 922-3899
ted.stevenson@alston.com
brady.cox@alston.com

Erik John Carlson (CA 265167)
Caleb J. Bean (CA 299751)
Srishti Ghosh (CA 354393)
ALSTON & BIRD LLP
350 South Grand Avenue, 51st Floor

Los Angeles, CA 90071
Telephone: (213) 576-1000
Facsimile: (213) 576-1100
erik.carlson@alston.com
caleb.bean@alston.com
maddie.ghosh@alston.com

Natalie C. Clayton (NY 4409538)
ALSTON & BIRD LLP
90 Park Avenue, 15th Floor
New York, NY 10016
Telephone: (212) 210-9400
Facsimile: (212) 210-9444
natalie.clayton@alston.com

David Greenbaum (NY 2947455)
GREENBAUM LAW LLC
210 Allison Court
Englewood, NJ 07631
david@greenbaum.law

VBA-Samsung@alston.com

Attorneys for VB Assets LLC

Date: April 21, 2025

/Lori A. Gordon/

Lori A. Gordon (Reg. No. 50,633)
Goodwin Procter LLP
1900 N Street, N.W.
Washington, D.C. 20036
Phone: (202) 346 4435
Gordon-ptab@goodwinlaw.com

Lead Counsel for Petitioner

APPENDIX A – LIST OF CLAIMS

[1P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[1A] receiving, at the one or more physical processors, a first natural language utterance;

[1B] providing, by the one or more physical processors, a response to the first natural language utterance;

[1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[1D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance;

[1E] determining, by the one or more physical processors, domain information for the one or more recognized words based on the first natural language utterance;

[1F] processing, by the one or more physical processors, based on the domain information, the one or more recognized words to determine an interpretation of the second natural language utterance, wherein processing the one or more recognized words comprises:

[1G] providing the one or more recognized words to a first domain agent associated with a first domain and a second domain agent associated with a second domain;

[1H] obtaining a first interpretation of the second natural language utterance from the first domain agent;

[1I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and

[1J] determining the interpretation based on one or more of the first interpretation or the second interpretation;

[1K] determining, by the one or more physical processors, promotional content based on the interpretation; and

[1L] presenting, by the one or more physical processors, the promotional content to a user.

[2A] The method of claim 1, wherein the first natural language utterance and the second natural language utterance are received during an interactive session between the user and the computer system, the method further comprising:

[2B] determining, by the one or more physical processors, an activity relating to one or more of the first natural language utterance or the second natural language utterance,

[2C] wherein determining the promotional content comprises determining the promotional content based on the activity.

[3A] The method of claim 1, further comprising: determining, by the one or more physical processors, context information for the one or more recognized words based on the first natural language utterance,

[3B] wherein the interpretation is determined further based on the context information.

[4] The method of claim 1, wherein the first interpretation is determined by the first domain agent based on the first domain, and wherein the second

interpretation is determined by the second domain agent based on the second domain.

[5A] The method of claim 1, further comprising: associating, by the one or more physical processors, the first interpretation with a first interpretation score; and

[5B] associating, by the one or more physical processors, the second interpretation with a second interpretation score,

[5C] wherein determining the interpretation comprises determining the first interpretation as the interpretation based on the first interpretation score and the second interpretation score.

[6] The method of claim 5, wherein the first interpretation is determined as the interpretation based on the first interpretation score being greater than the second interpretation score.

[7A] The method of claim 1, further comprising: receiving, at the one or more physical processors, a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance; and

[7B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content.

[8] The method of claim 7, wherein the first natural language utterance, the second natural language utterance, and the third natural language utterance are received during an interactive session between the user and the computer system.

[9] The method of claim 8, wherein the third natural language utterance is received after the second natural language utterance.

[10A] The method of claim 7, further comprising: performing, by the one or more physical processors, speech recognition to recognize one or more words of the third natural language utterance, wherein the one or more recognized words of the third natural language utterance include one or more reference words;

[10B] determining, by the one or more physical processors, that the one or more reference words refer to an item associated with the promotional content; and
initiating, by the one or more physical processors, a purchase transaction associated with the item based on the determination that the one or more reference words refer to the item.

[11] The method of claim 10, wherein the item includes a purchasable good.

[12] The method of claim 7, further comprising: initiating, by the one or more physical processors, based on the interpretation of the third natural language utterance, a purchase transaction associated with the promotional content.

[13A] The method of claim 1, wherein the first natural language utterance and the second natural language utterance are received during an interactive session between the user and the computer system, the method further comprising:

[13B] receiving, at the one or more physical processors, a third natural language utterance during the interactive session; and

[13C] processing, by the one or more physical processors, based on the third natural language utterance, the one or more recognized words to determine a reinterpretation of the second natural language utterance.

[14] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the reinterpretation.

[15A] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising:

[15B] determining, by the one or more physical processors, other promotional content based on the reinterpretation; and

[15C] presenting, by the one or more physical processors, the other promotional content to the user.

[16A] The method of claim 1, further comprising: identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance,

[16B] wherein determining the promotional content comprises obtaining the promotional content based on a determination that the promotional content relates to the one or more requests.

[17A] The method of claim 16, further comprising: determining, by the one or more physical processors, one or more activities in the one or more requests; and

[17B] performing, by the one or more physical processors, the one or more activities,

[17C] wherein determining the promotional content comprises determining the promotional content based on the one or more activities.

[18] The method of claim 17, wherein determining the one or more activities comprises determining a call to be initiated for the user in response to the one or more requests, wherein performing the one or more activities comprises initiating the call for the user, and wherein determining the promotional content comprises determining the promotional content based on the call.

[19] The method of claim 17, wherein determining the one or more activities comprises determining a navigation route to be calculated for the user in response to the one or more requests, wherein performing the one or more activities comprises calculating the navigation route for the user, and wherein determining the promotional content comprises determining the promotional content based on the navigation route.

[20] The method of claim 17, wherein determining the one or more activities comprise determining a media item to be presented in response to the one or more requests, wherein performing the one or more activities comprises presenting the media item to the user, and wherein determining the promotional content comprises determining the promotional content based on the media item, the promotional content being different than the media item.

[21] The method of claim 17, wherein determining the one or more activities comprises determining a reservation to be made on behalf of the user in response to the one or more requests, and wherein determining the promotional content comprises determining the promotional content based on information relating to the reservation.

[22] The method of claim 17, wherein determining the one or more activities comprises determining a first reservation to be made on behalf of the user in response to the one or more requests, and wherein determining the promotional content comprises determining the promotional content based on information relating to a second reservation that is an alternative to the first reservation.

[23A] The method of claim 17, further comprising: receiving, at the one or more physical processors, a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance;

[23B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content; and

[23C] initiating, by the one or more physical processors, based on the interpretation of the third natural language utterance, a purchase transaction associated with the promotional content.

[24A] The method of claim 16, further comprising: determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous;

[24B] monitoring, by the one or more physical processors, interaction of the user with the promotional content; and

[24C] interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction.

[25A] The method of claim 16, wherein the one or more requests include a first request associated with a first application and a second request associated with a second application different than the first application, and

[25B] wherein the promotional content relates to one or more of the first request or the second request.

[26] The method of claim 25, wherein a first application type of the first application includes one or more of a navigation application, a music application, a commerce application, or a calling application, and wherein the second application is of an application type different than the first application type.

[27A] The method of claim 16, wherein the one or more requests include a first request to be processed by a first device associated with the user and a second request to be processed by a second device associated with the user, and

[27B] wherein the promotional content relates to one or more of the first request or the second request.

[28] The method of claim 27, wherein a first device type of the first device includes one or more of a mobile phone, a navigation device, or a media player device, and wherein the second device is of a device type different than the first device type.

[29] The method of claim 27, wherein the first device and the second device are operably independent of one another.

[30A] The method of claim 1, further comprising: obtaining, by the one or more physical processors, user profile information associated with the user, wherein the user profile information specifies prior user interactions with items; and

[30B] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information,

[30C] wherein determining the promotional content comprises determining a promotional item associated with one of the categories.

[31A] The method of claim 1, further comprising: obtaining, by the one or more physical processors, user profile information associated with the user, wherein the user profile information specifies prior user interactions with items;

[31B] identifying, by the one or more physical processors, one or more requests associated with the first natural language utterance or the second natural language utterance;

[31C] determining, by the one or more physical processors, one or more applications for processing the one or more requests; and

[31D] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information, wherein the categories relate to the one or more applications,

[31E] wherein determining the promotional content comprises determining a promotional item associated with one of the categories.

[32P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[32A] receive a first natural language utterance;

[32B] provide a response to the first natural language utterance;

[32C] receive a second natural language utterance relating to the first natural language utterance;

[32D] perform speech recognition to recognize one or more words of the second natural language utterance;

[32E] determine domain information for the one or more recognized words based on the first natural language utterance;

[32F] process, based on the domain information, the one or more recognized words to determine an interpretation of the second natural language utterance, wherein processing the one or more recognized words comprises:

[32G] providing the one or more recognized words to a first domain agent associated with a first domain and a second domain agent associated with a second domain;

[32H] obtaining a first interpretation of the second natural language utterance from the first domain agent;

[32I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and

[32J] determining the interpretation based on one or more of the first interpretation or the second interpretation;

[32K] determine promotional content based on the interpretation; and

[32L] present the promotional content to a user.

[33P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[33A] receiving, at the one or more physical processors, a first natural language utterance during an interactive session between a user and the computer system;

[33B] providing, by the one or more physical processors, a response to the first natural language utterance;

[33C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance during the interactive session;

[33D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance;

[33E] processing, by the one or more physical processors, based on the first natural language utterance, the one or more recognized words to determine an interpretation of the second natural language utterance;

[33F] receiving, at the one or more physical processors, a third natural language utterance during the interactive session;

[33G] processing, by the one or more physical processors, based on the third natural language utterance, the one or more recognized words to determine a reinterpretation of the second natural language utterance;

[33H] determining, by the one or more physical processors, promotional content based on one or more of the interpretation or the reinterpretation; and

[33I] presenting, by the one or more physical processors, the promotional content to the user.

[34] The method of claim 33, wherein determining the promotional content comprises determining the promotional content based on the reinterpretation.

[35A] The method of claim 33, wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising:

[35B] determining, by the one or more physical processors, other promotional content based on the reinterpretation; and

[35C] presenting, by the one or more physical processors, the other promotional content to the user.

[36P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[36A] receive a first natural language utterance during an interactive session between a user and the system;

[36C] provide a response to the first natural language utterance;

[36C] receive a second natural language utterance relating to the first natural language utterance during the interactive session;

[36D] perform speech recognition to recognize one or more words of the second natural language utterance;

[36E] process, based on the first natural language utterance, the one or more recognized words to determine an interpretation of the second natural language utterance;

[36F] receive a third natural language utterance during the interactive session;

[36G] process, based on the third natural language utterance, the one or more recognized words to determine a reinterpretation of the second natural language utterance;

[36H] determine promotional content based on one or more of the interpretation or the reinterpretation; and

[36I] present the promotional content to the user.

[37P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[37A] receiving, at the one or more physical processors, a first natural language utterance;

[37B] providing, by the one or more physical processors, a response to the first natural language utterance;

[37C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[37D] identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance;

[37E] determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous;

[37F] determining, by the one or more physical processors, promotional content that relates to the one or more requests;

[37G] presenting, by the one or more physical processors, the promotional content to a user;

[37H] monitoring, by the one or more physical processors, interaction of the user with the promotional content; and

[37I] interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction.

[38P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more

physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

- [38A] receive a first natural language utterance;
- [38B] provide a response to the first natural language utterance;
- [38C] receive a second natural language utterance relating to the first natural language utterance;
- [38D] identify one or more requests associated with the second natural language utterance;
- [38E] determine that at least one request of the one or more requests is incomplete or ambiguous;
- [38F] determine promotional content that relates to the one or more requests;
- [38G] present the promotional content to a user;
- [38H] monitor interaction of the user with the promotional content; and
- [38I] interpret the at least one incomplete or ambiguous request based on the interaction.

[39P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

- [39A] receiving, at the one or more physical processors, a first natural language utterance;
- [39B] providing, by the one or more physical processors, a response to the first natural language utterance;
- [39C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[39D] identifying, by the one or more physical processors, requests associated with the second natural language utterance, wherein the requests include a first request associated with a first application and a second request associated with a second application different than the first application;

[39E] determining, by the one or more physical processors, promotional content that relates to one or more of the first request or the second request; and

[39F] presenting, by the one or more physical processors, the promotional content to a user.

[40] The method of 39, wherein a first application type of the first application includes one or more of a navigation application, a music application, a commerce application, or a calling application, and wherein the second application is of an application type different than the first application type.

[41A] The method of claim 39, further comprising: receiving, at the one or more physical processors, a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance; and

[41B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content.

[42] The method of claim 41, further comprising: initiating, by the one or more physical processors, based on the interpretation of the third natural language utterance, a purchase transaction associated with the promotional content.

[43P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more

physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[43A] receive a first natural language utterance;

[43B] provide a response to the first natural language utterance;

[43C] receive a second natural language utterance relating to the first natural language utterance;

[43D] identify requests associated with the second natural language utterance, wherein the requests include a first request associated with a first application and a second request associated with a second application different than the first application;

[43E] determine promotional content that relates to one or more of the first request or the second request; and

[43F] present the promotional content to a user.

[44P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[44A] receiving, at the one or more physical processors, a first natural language utterance;

[44B] providing, by the one or more physical processors, a response to the first natural language utterance;

[44C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[44D] identifying, by the one or more physical processors, requests associated with the second natural language utterance, wherein the requests include

a first request to be processed by a first device associated with a user and a second request to be processed by a second device associated with the user;

[44E] determining, by the one or more physical processors, promotional content that relates to one or more of the first request or the second request; and

[44F] presenting, by the one or more physical processors, the promotional content to the user.

[45] The method of claim 44, wherein a first device type of the first device includes one or more of a mobile phone, a navigation device, or a media player device, and wherein the second device is of a device type different than the first device type.

[46] The method of claim 44, wherein the first device and the second device are operably independent of one another.

[47A] The method of claim 44, further comprising: receiving, at the one or more physical processors, a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance; and

[47B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content.

[48] The method of claim 47, further comprising: initiating, by the one or more physical processors, based on the interpretation of the third natural language utterance, a purchase transaction associated with the promotional content.

[49P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more

physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[49A] receive a first natural language utterance;

[49B] provide a response to the first natural language utterance;

[49C] receive a second natural language utterance relating to the first natural language utterance;

[49D] identify requests associated with the second natural language utterance wherein the requests include a first request to be processed by a first device associated with a user and a second request to be processed by a second device associated with the user;

[49E] determine promotional content that relates to one or more of the first request or the second request; and

[49F] present the promotional content to the user.

[50P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[50A] receiving, at the one or more physical processors, a first natural language utterance;

[50B] providing, by the one or more physical processors, a response to the first natural language utterance;

[50C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[50D] obtaining, by the one or more physical processors, user profile information associated with a user, wherein the user profile information specifies prior user interactions with items;

[50E] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information;

[50F] determining, by the one or more physical processors, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and

[50G] presenting, by the one or more physical processors, the promotional content to the user.

[51A] The method of claim 50, further comprising: receiving, at the one or more physical processors, a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance; and

[51B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content.

[52] The system of claim 51 wherein the plurality of entries in the one or more dictionary and phrase tables are further dynamically updated based on one or more dynamic fuzzy set possibilities or prior probabilities derived from the history of the current dialog and the prior dialogs.

[53P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[53A] receive a first natural language utterance;

[53B] provide a response to the first natural language utterance;

[53C] receive a second natural language utterance relating to the first natural language utterance;

[53D] obtain user profile information associated with a user, wherein the user profile information specifies prior user interactions with items;

[53E] identify categories of items based on the prior user interactions specified by the user profile information;

[53F] determine, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and

[53G] present the promotional content to the user.

[54P] A computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

[54A] receiving, at the one or more physical processors, a first natural language utterance;

[54B] providing, by the one or more physical processors, a response to the first natural language utterance;

[54C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;

[54D] obtaining, by the one or more physical processors, user profile information associated with a user, wherein the user profile information specifies prior user interactions with items;

[54E] identifying, by the one or more physical processors, one or more requests associated with one or more of the first natural language utterance or the second natural language utterance;

[54F] determining, by the one or more physical processors, one or more applications for processing the one or more requests; and

[54G] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information, wherein the categories relate to the one or more applications;

[54H] determining, by the one or more physical processors, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and

[54I] presenting, by the one or more physical processors, the promotional content to the user.

[55P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:

[55A] receive a first natural language utterance;

[55B] provide a response to the first natural language utterance;

[55C] receive a second natural language utterance relating to the first natural language utterance;

[55D] obtain user profile information associated with a user, wherein the user profile information specifies prior user interactions with items;

[55E] identify one or more requests associated with one or more of the first natural language utterance or the second natural language utterance;

[55F] determine one or more applications for processing the one or more requests; and

[55G] identify categories of items based on the prior user interactions specified by the user profile information, wherein the categories relate to the one or more applications;

[55H] determine, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and

[55I] present the promotional content to the user.