**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

---

# DECLARATION OF STUART LIPOFF IN SUPPORT OF 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

**Page**

## TABLE OF CONTENTS


I. Qualifications ..... 3

II. Understanding of Relevant Legal Principles ..... 9

III. Technical Background ..... 11

A. Language and Dialog ..... 11

B. Interactive/Spoken Dialog Systems ..... 13

1. Speech Recognition Component ..... 17

2. Language Understanding ..... 19

3. Dialog Management Component ..... 21

4. Response Generator and Speech Synthesizer ..... 22

C. Using Domain Agents to Improve Spoken Dialogue Systems ..... 23

D. Intent-Based Advertising ..... 26

E. Web-Based Advertising Tools ..... 28

1. Early Web-Advertising Systems ..... 29

2. User Modelling Approaches ..... 32

3. Using Analytics to Improve Models ..... 36

IV. '536 Patent ..... 38

A. Overview of the '536 Patent ..... 38

B. Level of Ordinary Skill in the Art ..... 49

C. Prosecution and Litigation History ..... 50

1. U.S. Prosecution Before the U.S. Patent & Trademark Office ..... 50

2. Inter Partes Review before the Patent Trial and Appeal Board ..... 52

3. Proceedings in the U.S. District Court ..... 56

V. GROUND 1: THE COMBINATION OF O'NEILL AND JONES RENDERS OBVIOUS CLAIMS 1-12, 16-23, 30, 32 AND 50-55 ..... 56

A. Overview of the Combination ..... 56

**TABLE OF CONTENTS**
(continued)



1. O'Neill .....57
2. Jones .....62
3. Motivation to Combine .....66

B. Independent Claims 1 and 32 .....70
1. Preamble [1P] / [32P] .....73
2. Receiving Natural Language Request [1A] / [32A] .....77
3. Providing a Response [1B] / [32B] .....79
4. Receiving a Second Natural Language Utterance [1C] / [32C] .....82
5. Performing Speech Recognition on One or More Words in Second Natural Language Utterance [1D] / [32D] .....84
6. Determining Domain Information Based on First Natural Language Utterance [1E] / [32E] .....88
7. Interpreting the Second Natural Language Utterance Based on the Domain Information [1F] / [32F] .....92
8. Providing Recognized Words to First and Second Domain Agents [1G] / [32G] .....93
9. Obtaining Interpretations of Second Natural Language Utterance from First and Second Domain Agents [1H] / [32H] and [1I] / [32I] .....97
10. Determining Interpretation [1J] / [32J] .....100
11. Determining Promotional Content [1K] / [32K] .....104
12. Presenting the Promotional Content [1L] / [32L] .....105

C. Dependent Claim 2 .....106
1. Utterances Received During Interactive Session [2A] .....106
2. Determining an Activity Relating to One or More of the Natural Language Utterances [2B] .....107
3. Determining the Promotional Content Based on the Activity [2C] .....110

D. Dependent Claim 3 .....112

**TABLE OF CONTENTS**
(continued)



1. Determining Context Information Based on the First Natural Language Utterance [3A] ..... 112
2. Interpretation of the Second Natural Language Utterance Based on Context Information [3B] ..... 116

E. Dependent Claim 4 ..... 119

F. Dependent Claim 5 ..... 120

1. Associating the Interpretations With Interpretation Scores [5A] and [5B] ..... 120
2. Determining the First Interpretation Based on the First and Second Interpretation Score [5C] ..... 122

G. Dependent Claim 6 ..... 123

H. Dependent Claims 7 and 51 ..... 124

1. Receiving Third Natural Language Utterance [7A] / [51A] ..... 125
2. Determining an Interpretation of the Third Natural Language Utterance Based on the Promotional Content [7B] / [51B] ..... 128

I. Dependent Claim 8 ..... 131

J. Dependent Claim 9 ..... 133

K. Dependent Claim 10 ..... 134

1. Performing Speech Recognition on Third Natural Language Utterance [10A] ..... 134
2. Determining That One or More Reference Words Refer to an Item Associated with the Promotional Content [10B] ..... 137
3. Initiating a Purchase Transaction [10C] ..... 139

L. Dependent Claim 11 ..... 141

M. Dependents Claims 12 and 52 ..... 142

N. Dependent Claim 16 ..... 145

1. Identifying One or More Requests Associated With Second Natural Language Utterance [16A].....145
2. Obtaining Promotional Content Related to Request [16B] ...148

O. Dependent Claim 17.....149
1. Determining One or More Activities in the One or More Requests [17A].....149
2. Performing the One or More Activities [17B].....150
3. Determining Promotional Content Based on Activity [17C] .....151

P. Dependent Claim 18.....152

Q. Dependent Claim 19.....154

R. Dependent Claim 20.....157

S. Dependent Claim 21.....159
1. Determining a Reservation For the User in Response to One or More Requests [21A].....159
2. Determining Promotional Content Based on Information Related to the Reservation [21B].....160

T. Dependent Claim 22.....161
1. Determining a First Reservation in Response to User Request [22A] .....161
2. Determining Promotional Content Based on Alternative Reservation [22B] .....164

U. Dependent Claim 23.....165
1. Receiving a Third Natural Language Utterance [23A].....165
2. Determining an Interpretation of Third Natural Language Utterance Based on Promotional Content [23B] .....166
3. Initiating a Purchase Transaction [23C] .....167

V. Dependent Claim 30.....168
1. Obtaining User Profile Information [30A] .....168

2. Identifying Categories of Items Based on Prior User Interactions [30B] .....172

3. Determining a Promotional Item Associated with One or More Categories [30C] .....177

W. Independent Claim 50 .....179

1. Preamble [50P].....179

2. Receiving a First Natural Language Utterance [50A] .....180

3. Providing a Response to the First Natural Language Utterance [50B].....181

4. Receiving a Second Natural Language Utterance [50C].....181

5. Obtaining User Profile Information [50D] .....182

6. Identifying Categories of Items Based on Prior User Interactions [50E].....183

7. Determining Promotional Content Associated with One of the Categories Based on the First and Second Natural Language Utterances [50F].....183

8. Presenting Promotional Content [50G].....187

X. Independent Claim 53 .....187

1. Preamble [53P].....187

2. Receive a First Natural Language Utterance [53A] .....189

3. Provide Response to First Natural Language Utterance [53B] .....189

4. Receive a Second Natural Language Utterance [53C] .....190

5. Obtaining User Profile Information [53D] .....190

6. Identify Categories of Items Based on Prior User Interactions [53E].....191

7. Determine Promotional Content Associated with One of the Categories [53F].....192

8. Present Promotional Content [53G].....193

Y. Independent Claim 54 .....193

1. Preamble [54P].....193
2. Receiving a First Natural Language Utterance [54A].....194
3. Providing a Response to the First Natural Language Utterance [54B].....195
4. Receiving a Second Natural Language Utterance [54C].....195
5. Obtaining User Profile Information [54D].....196
6. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [54E].....197
7. Determining Application for Processing One or More Requests [54F].....198
8. Identifying Categories of Items Based on Prior Use Interactions Related to One or More Applications [54G].....199
9. Determining Promotional Content Associated with One of the Categories [54H].....200
10. Presenting the Promotional Content [54I].....200

Z. Independent Claim 55.....201
1. Preamble [55P].....201
2. Receive a First Natural Language Utterance [55A].....202
3. Provide Response to First Natural Language Utterance [55B].....203
4. Receive a Second Natural Language Utterance [55C].....203
5. Obtain User Profile Information [55D].....204
6. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [55E].....205
7. Determine Application for Processing One or More Requests [55F].....206
8. Identify Categories of Items Based on Prior Use Interactions Related to One or More Applications [55G].....206
9. Determine Promotional Content Associated with One of the Categories [55H].....207

10. Present Promotional Content [55I] .....208

VI. GROUND 2: THE COMBINATION OF O'NEILL, JONES AND FRANCO RENDERS OBVIOUS CLAIMS 13-15, 24 AND 33-38 .....209

A. Overview of the Combination .....209

1. Franco .....209

2. Motivation to Combine .....211

B. Dependent Claim 13 .....212

1. Interactive Session [13A] .....212

2. Receiving a Third Natural Language Utterance During the Interactive Session [13B] .....213

3. Processing the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [13C] .....214

C. Dependent Claim 14 .....223

D. Dependent Claim 15 .....225

1. Determining the Promotional Content Based on the Interpretation [15A] .....225

2. Determining Other Promotional Content Based on the Reinterpretation [15B] .....226

3. Presenting the Promotional Content to the User [15C] .....228

E. Dependent Claim 24 .....228

1. Determining at Least One Request is Incomplete or Ambiguous [24A] .....228

2. Monitoring Interactions of the User with the Promotional Content [24B] .....232

3. Interpreting the Incomplete or Ambiguous Request Based on the Interaction [24C] .....236

F. Independent Claim 33 .....238

1. Preamble [33P] .....238

2. Receiving a First Natural Language Utterance [33A] .....239

3. Providing a Response to the First Natural Language Utterance [33B] ..... 240
4. Receiving a Second Natural Language Utterance [33C] ..... 241
5. Performing Speech Recognition to Recognize Words of Second Natural Language Utterance [33D] ..... 242
6. Determining an Interpretation of the Second Natural Language Utterance [33E] ..... 243
7. Receiving a Third Natural Language Utterance [33F] ..... 244
8. Processing the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [33G] ..... 244
9. Determining Promotional Content [33H] ..... 245
10. Receiving a Second Natural Language Utterance [33I] ..... 246

G. Dependent Claim 34 ..... 247

H. Dependent Claim 35 ..... 248

1. Determining the Promotional Content Based on the Interpretation [35A] ..... 248
2. Determining Other Promotional Content Based on the Interpretation [35B] ..... 248
3. Presenting the Promotional Content to the User [35C] ..... 249

I. Independent Claim 36 ..... 250

1. Preamble [36P] ..... 250
2. Receive a First Natural Language Utterance [36A] ..... 251
3. Provide Response to First Natural Language Utterance [36B] ..... 251
4. Receive a Second Natural Language Utterance [36C] ..... 252
5. Perform Speech Recognition to Recognize One or More Words in Second Natural Utterance [36D] ..... 253

6. Process One or More Recognized Words to Determine Interpretation of Second Natural Language Utterance [36E].....253
7. Receive a Third Natural Language Utterance [36F].....254
8. Process the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [36G].....255
9. Determine Promotional Content Based on One or More of the Interpretation or the Reinterpretation [36H].....256
10. Present the Promotional Content to the User [36I].....256

J. Independent Claim 37 .....257
1. Preamble [37P].....257
2. Receiving a First Natural Language Utterance [37A] .....258
3. Providing a Response to the First Natural Language Utterance [37B].....259
4. Receiving a Second Natural Language Utterance [37C].....259
5. Identifying a Request [37D].....260
6. Incomplete or Ambiguous Request [37E] .....261
7. Determining Promotional Content Related to Request [37F].....261
8. Present Promotional Content [37G].....262
9. Identifying a Request [37H].....263
10. Incomplete or Ambiguous Request Based on the Interaction [37I] .....264

K. Independent Claim 38 .....265
1. Preamble [38P].....265
2. Receive a First Natural Language Utterance [38A] .....266
3. Provide Response to First Natural Language Utterance [38B] .....266
4. Receive a Second Natural Language Utterance [36C] .....267

5. Identify One or More Requests Associated with the Second Natural Language Utterance [38D] ..... 267

6. Determine at Least One Request is Incomplete or Ambiguous [38E] ..... 268

7. Determine Promotional Content Relating to Request [38F] ..... 269

8. Present the Promotional Content to the User [38G] ..... 269

9. Monitor Interactions of the User with the Promotional Content [38H] ..... 270

10. Interpret the Incomplete or Ambiguous Request Based on the Interaction [38I] ..... 271

I. GROUND 3: THE COMBINATION OF O'NEILL, JONES AND WANG RENDERS OBVIOUS CLAIMS 25-26, 31 AND 39-43 ..... 271

A. Combination Overview ..... 272

1. Wang ..... 272

2. Motivation to Combine ..... 273

B. Dependent Claim 25 ..... 274

1. First Request Associated with First Application and Second Request Associated with Second Application [25A] ..... 274

2. Promotional Content Related to One or More of the First or Second Request [25B] ..... 277

C. Dependent Claims 26 and 40 ..... 278

D. Dependent Claim 31 ..... 279

1. Obtaining User Profile Information [31A] ..... 279

2. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [31B] ..... 279

3. Determining One or More Application for Processing One or More Requests [31C] ..... 281

4. Identifying Categories of Items Based on Prior Use Interactions Related to One or More Applications [31D] ..... 282

5. Determining Promotional Item Associated with One or More Categories [31E] ..... 284

E. Independent Claim 39 ..... 285

1. Preamble [39P] ..... 285

2. Receiving a First Natural Language Utterance [39A] ..... 286

3. Providing a Response to the First Natural Language Utterance [39B] ..... 287

4. Receiving a Second Natural Language Utterance [39C] ..... 287

5. Identifying Requests Associated With Applications [39D] / Determining Promotional Content Relating to Request [39E] ..... 288

6. Presenting Promotional Content [39F] ..... 289

F. Dependent Claim 41 ..... 290

G. Dependent Claim 42 ..... 291

H. Independent Claim 43 ..... 292

1. Preamble [43P] ..... 292

2. Provide Response to First Natural Language Utterance [43B] ..... 293

3. Receive a Second Natural Language Utterance [43C] ..... 293

4. Identify Requests Associated With Applications [43D] ..... 294

5. Determine Promotional Content Relating to Request [43E] ..... 295

6. Present Promotional Content [43F] ..... 296

II. GROUND 4: THE COMBINATION OF O'NEILL, JONES AND HOWARD RENDERS OBVIOUS CLAIMS 27-29 AND 44-49 ..... 296

A. Overview of the Combination ..... 297

1. Howard ..... 297

2. Motivation to Combine ..... 299

B. Dependent Claim 27 ..... 302

1. First Request Processed by a First Device and Second Request Processed by Second Device [27A] ..... 302
2. Promotional Content Relates to One or More Requests [27B] ..... 305

C. Dependent Claims 28 and 45 ..... 306

D. Dependent Claims 29 and 46 ..... 307

E. Independent Claim 44 ..... 308
1. Preamble [44P] ..... 308
2. Receiving Natural Language Utterance [44A] ..... 309
3. Providing a Response [44B] ..... 310
4. Receiving a Second Natural Language Utterance [44C] ..... 311
5. Identifying Requests Associated with the Second Utterance, Wherein First Request is Processed by a First Device and Second Request is Processed by a Second Device [44D] ..... 311
6. Determining Promotional Content Related to One or More Requests [44E] ..... 312
7. Presenting the Promotional Content [44F] ..... 313

F. Dependent Claim 47 ..... 314
1. Receiving Third Natural Language Utterance [47A] ..... 314
2. Determining an Interpretation of the Third Natural Language Utterance Based on the Promotional Content [47B] ..... 314

G. Dependent Claim 48 ..... 315

H. Independent Claim 49 ..... 316
1. Preamble [49P] ..... 316
2. Receive a First Natural Language Utterance [49A] ..... 317
3. Provide Response to First Natural Language Utterance [49B] ..... 318
4. Receive a Second Natural Language Utterance [49C] ..... 318

5. Identify Requests Associated with the Second Utterance, Wherein First Request is Processed by a First Device and Second Request is Processed by a Second Device [49D].....319

6. Determine Promotional Content Related to One or More Requests [49E] .....320

7. Presenting the Promotional Content [49F] .....320

III. CONCLUSION.....321

I, Stuart Lipoff, declare as follows:

1. I have been retained by Goodwin Procter LLP on behalf of Petitioners Samsung Electronics Co. Ltd. and Samsung Electronics America, Inc. ("Petitioners") to provide this Declaration concerning technical subject matter relevant to the petition for *Inter Partes* Review ("Petition") of U.S. Patent 8,886,536 ("the '536 patent"; EX-1001).

2. I am over 18 years of age. I have personal knowledge of the facts stated in this Declaration and could testify competently to them if asked to do so.

3. I have been asked to provide my technical opinions regarding how a person of ordinary skill in the art would have understood the claims of the '536 patent at the time of the alleged invention, which I have been asked to assume is the 2007 timeframe. For purposes of whether the teachings of the prior art render the claims of the '536 patent obvious, I have been asked to assume the date of February 6, 2007, which I understand to be the filing date of the earliest filed application that lead to the '536 patent. I have also been asked to provide my technical opinions on how concepts in the '536 patent specification relate to claim limitations of the '536 patent.

4. In reaching the opinions provided herein, I have considered the '536 patent, its prosecution history, and the references cited in my Declaration and the

Appendix. I have also drawn on my own education, training, research, knowledge, and personal and professional experience.

5. In general, I have been asked to cite to the specification of a patent or patent publication using the following formats: (Patent, Col:Line Number(s)) or (Patent, Paragraph Number(s)). For example, the citation (’536 patent, 1:1-10) points to the ’536 patent specification at column 1, lines 1-10. Also, for convenience, I have been asked to use italics to denote limitations from the challenged claims. Unless otherwise noted, all emphasis is added.

6. I am being compensated for my time at my standard consulting rate. I am also being reimbursed for expenses that I incur during the course of this work. My compensation is not contingent upon the results of my study and analysis, the substance of my opinions, or the outcome of any proceeding involving the ’536 patent. I have no financial interest in the outcome of this matter or in any litigation involving the ’536 patent.

7. I understand that another party, Amazon.com, Inc., Amazon.com LLC, Amazon Web Services, Inc., A2Z Development Center, Inc. d/b/a Lab126, Rawles LLC, AMZN Mobile LLC, AMZN Mobile 2 LLC, Amazon.com Services, Inc. f/k/a Amazon Fulfillment Services, Inc., and Amazon.com Services LLC (formerly Amazon Digital Services LLC) (collectively “Amazon”), previously challenged the ’536 patent in two *inter partes* review proceedings at *Amazon.com, Inc. et al. v. VB*

*Assets, LLC*, IPR2020-01374 (P.T.A.B., Apr. 19, 2021) and *Amazon.com, Inc. et al. v. VB Assets, LLC*, IPR2020-01377 (P.T.A.B., Feb. 4, 2021) (collectively, the "Amazon-VBA-IPRs"), which I understand were denied instituted by the Patent Trial and Appeal Board. I refer in my declaration below to some of the positions and issues raised in that proceeding, including as it pertains to claim construction of certain claim terms in the '536 patent.

## I. Qualifications

8. I believe I am well qualified to render useful opinions on this matter. I will briefly summarize my knowledge, training, and experience here. A more detailed summary of my background, education, experience, and publications is set forth in my curriculum vitae (CV), which is submitted as EX-1004.

9. I am currently president of IP Action Partners Inc., a consulting practice that serves the telecommunications, information technology, media, electronics, and e-business industries.

10. I earned a Bachelor of Science degree in Electrical Engineering in 1968 and a second Bachelor of Science degree in Engineering Physics in 1969, both from Lehigh University. I earned a Master of Science degree in Electrical Engineering from Northeastern University in 1974, and then a Master of Business Administration degree from Suffolk University in 1983.

11. I hold a Federal Communications Commission ("FCC") General Radiotelephone Operator License. I also hold a Certificate in Data Processing from the Institute for the Certification of Computing Professionals ("ICCP"), which is supported by the Association for Computing Machinery ("ACM").

12. I am also a registered professional engineer (PE) in the Commonwealth of Massachusetts and in the State of Nevada.

13. I am a fellow of the Institute of Electrical and Electronics Engineers ("IEEE") Consumer Electronics, Communications, Computer, Circuits, and Vehicular Technology Groups. I have been a member of the IEEE Consumer Electronics Society National Board of Governors (formerly known as the Administrative Committee) since 1981, and I was Boston Chapter Chairman of the IEEE Vehicular Technology Society from 1974 to 1976. I served as the 1996-1997 President of the IEEE Consumer Electronics Society, and from 1999 to 2018 I served as Chairman of the Society's Technical Activities and Standards Committee and as Vice President of Publications for the Society. From 2018 to 2023 I served as Vice President of Standards and Industry Activities for the Society and currently serve on the Board of Governors as the Historian for the Society. I have also served as an Mazura Ibuka Customer Technology Award committee member for the IEEE's Award in the field of consumer electronics.

14. I have prepared and presented numerous papers at the IEEE and at other professional meetings. For example, in fall 2000, I served as general program chair for IEEE's Vehicular Technology Conference on advanced wireless communication technology. I have organized sessions at The International Conference on Consumer Electronics, and I was the 1984 program chairman. I conducted an eight-week IEEE-sponsored short course on Fiber Optics System Design. I received IEEE's Centennial Medal in 1984, and I received IEEE's Millennium Medal in 2000.

15. As Vice President and Standards Group Chairman for the Association of Computer Users ("ACU") from 1980 to 1983, I served as the ACU representative to the ANSI X3 Standards Group. From 1976 to 1978, I served as Chairman of the task group on user rule compliance for the FCC's Citizens Advisory Committee on Citizen's Band radio.

16. Over the last 25 years, I have been a member of the Society of Cable Television Engineers, the Association for Computing Machinery, and the Society of Motion Picture and Television Engineers. From 2001 to 2004, I served as a member of the USA advisory board to the National Science Museum of Israel. In 1998, I presented a short course on international product development strategies as a faculty member for Technion Institute of Management in Israel. From 2001 to 2003, I served as a member of the board or directors of the Massachusetts Future Problem Solving Program.

17. I am a named inventor on seven United States patents and have several publications on data communications in publications, including Electronics Design, Microwaves, EDN, the Proceedings of the Frequency Control Symposium, Optical Spectra, and IEEE publications.

18. During my over fifty-year professional career, I have been heavily engaged in the study, analysis, evaluation, design, and implementation of products and technology associated with consumer electronics and electronic appliances. A particular focus of my professional activities has been improving the man-machine interface including voice, speech, and speaker recognition for man-machine interactions. I also have extensive experience in foundation technologies and the applications supporting speech signal processing.

19. For approximately three years, from 1969 to 1972, I served as Project Engineer for Motorola's Communications Division, where I had project design responsibilities for paging and wireless communication products. Projects that I worked on while employed at Motorola included work on paging systems that included digital voice storage, voice compression, and voice synthesis. I also worked on projects that interfaced wireless data communications terminals to public safety computer systems for mobile data retrieval and data entry.

20. For approximately four years, from 1972 to 1976, I served as Section Manager for Bell & Howell Communications Company, where I also had project

design responsibilities for paging and wireless communication products. I supported projects developing covert audio intelligence systems that recognized speech and activated digital voice compression recording systems. I also led projects for voice-based radio paging systems that recorded speech input, processed the speech to remove silence, processed the speech to digitally compress the speech, and store and forward the speech upon demand from dual tone multi-frequency or computer keyboard retrieval from the servers.

21. For 25 years from 1976 to 2001, I worked for Arthur D. Little, Inc. ("ADL"), where I became the Vice President and Director of Communications, Information Technology, and Electronics (CIE) and served in that role for 10 years, from 1991 to 2001. At ADL, I was responsible for the firm's global CIE practice in laboratory-based contract engineering, product development, and technology-based consulting. I was also involved in multiple pioneering efforts to identify and explore customer-to-business and business-to-business electronic commerce and transactions information processing opportunities, or "e-commerce". These projects involved technology assessment and analysis as well as developing architectures and systems to support multiple applications. The technologies typically involved an information retrieval component.

22. While at ADL, I worked on several projects involving the combination of voice interfaces (including speech recognition and voice audio output) and

information retrieval. Over the course of three years in the early-1990s, I worked on a project for Bolt Beranek and Newman (BB&N), where I evaluated and benchmarked technology for a voice input/output application that allowed end users (*e.g.*, travel agents) to use speech inputs to interact with airline reservation databases to retrieve information about travel reservation options, where the results were returned to the user in an audible message. This system included a natural language front-end speech-interface module with speech recognition that used pre- defined recognition grammars to convert the end user's speech into structured commands supported by an airline reservation system. As another example, over the course of three years in the mid-1990s, I also worked on a project for Texas Instruments that applied a speech–recognition interface for a variety of applications that retrieved information from database servers.

23. Other projects that I worked on at various points in my twenty five years at ADL and afterwards involving speech recognition technologies including systems for Sprint, Rockwell Semiconductor, and the United States Postal Service ("USPS"), detailed below.

24. Over the course of three years in the early 1990s, I worked on a voice–interface project developing spoken digit telephone number recognition and voiceprint matching for Sprint's long distance alternative access telephone services.

25. Over the course of a year in the late 1980s, I worked on a voice interface project evaluating the processing power needed to perform various voice recognition applications by Rockwell Semiconductor's signal processing technology.

26. Over the course of 15 years starting in the early 1980s, I worked on a project for the USPS, where we developed a real-time automated postal teller system that served as an interface between end-users and the USPS's information systems. This system included voice prompts for the vision impaired.

## II. Understanding of Relevant Legal Principles

27. I am not a lawyer, and I will not provide any legal opinions. Although I am not a lawyer, I have been advised certain legal standards are to be applied by technical experts in forming opinions regarding the meaning and validity of patent claims.

28. I understand that a patent claim is invalid if it is anticipated or obvious in view of the prior art, and that a claim can be unpatentable even if all of the requirements of the claim cannot be found in a single prior-art reference. I further understand that invalidity of a claim requires that the claim be anticipated or obvious from the perspective of a person of ordinary skill in the art at the time the invention was made.

29. I have been informed that a patent claim is invalid if it would have been obvious to a person of ordinary skill in the art. In analyzing the obviousness of a

claim, I understand the following factors may be taken into account: (1) the scope and content of the prior art; (2) the differences between the prior art and the claims; (3) the level of ordinary skill in the art; and (4) any so called "secondary considerations" of non-obviousness, if they are present. I am not aware of any evidence of secondary considerations of non-obviousness relevant to the '536 patent. I reserve the right to supplement this Declaration if Patent Owner ("PO") introduces evidence of secondary considerations of non-obviousness.

30. I understand that to prove that prior art or a combination of prior art renders a patent obvious, it is necessary to:

> (1) identify the particular references that, singly or in combination, make the patent obvious;
> (2) specifically identify which elements of the patent claim appear in each of the asserted references; and
> (3) explain why a person of ordinary skill in the art would have combined the references, and how they would have done so, to create the inventions claimed in the patent. I further understand that exemplary rationales that may support a conclusion of obviousness include:
>
> - combining prior art elements according to known methods to yield predictable results;
> - simple substitution of one known element for another to obtain predictable results;
> - use of known technique(s) to improve similar devices (methods or products) in the same way;

- applying a known technique to a known device (method or product) ready for improvement to yield predictable results;
- "obvious to try" – choosing from a finite number of identified, predictable solutions with a reasonable expectation of success; known work in one field of endeavor may prompt variations of the work for use in either the same field or a different field based on design incentives or other market forces if the variations are predictable to a person of ordinary skill in the art; and
- some teaching, suggestion, or motivation in the prior art that would have led a person of ordinary skill in the art to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

31. I have been informed that, in considering obviousness, hindsight reasoning derived from the patent-at-issue may not be used.

## III. Technical Background

### A. Language and Dialog

32. Spoken language "is very different to written language" because "people typically do not follow rigid syntactic and morphological constraints in their utterances." (EX-1014, 49.) The "lack of written language formality in spontaneous spoken language makes linguistic analysis by machine both more difficult than, and different to, analysis of written language." (EX-1014, 49.)

33. As Huang explains "[i]t is natural to assume that the *sentence* is a clear and simple chunking unit for dialog, by analogy with written communication." (EX-

1015, 860 (emphasis in original).) But, because "sentences are artificially delimited in written text, researchers in dialog communication usually speak of the *utterance* as the basic unit." (EX-1015, 860 (emphasis in original).)

34. In the context of conversational speech, one or more utterances can comprise a "turn." A turn in conversation refers to the time one speaker holds the floor, often including multiple utterances—the smaller units of speech. Turns structure the flow of dialogue, beyond just the speech segments themselves. Turns are viewed as "building blocks for constructing a common task-oriented understanding among participants." (EX-1015, 861.)

35. Dialog is best understood not as a series of isolated turns but as structured discourse in which utterances are interdependent and contextually grounded. A dialog "consists of at least two turns, one by each speaker", and a coherent dialog "will exhibit discourse phenomena" whose interpretation "depends on the dialogue context." (EX-1016, 46.)

36. Two examples of how dialog presents these issues are anaphora and ellipsis. Anaphora is a linguistic phenomenon where a word or phrase refers back to another word or phrase mentioned earlier in the discourse. In dialog systems, anaphora resolution is essential for understanding references like pronouns. Ellipsis occurs when part of a sentence is omitted because it is understood from context. In

dialogue, users often leave out information that has already been established, and the system must infer the missing parts. (EX-1017, 62, 64.)

37. A key structuring device in dialogue is the adjacency pair—predictable turn sequences like question–answer or greeting–greeting—that capture mutual expectations and conversational obligations. (EX-1016, 65). These pairs are not merely formalities but reflect deeper discourse coherence, and they often aggregate into larger goal-driven patterns known as dialog games. (EX-1016, 65-68.) These games model the intentional structure of conversation and accommodate embedded exchanges, such as clarifications or corrections, within a larger purpose. (EX-1016, 65-68.)

38. Common terminology exists in the field of interactive/spoken dialog systems. For example, a "*lexicon* is a list of words, a *vocabulary*, annotated with syntactic (including morphological) and semantic features." (EX-1014, 49 (emphasis in original).) "*Grammars* describe how words may be combined into phrases and sentences." (EX-1014, 50 (emphasis in original).) "*Semantics* are abstract representations of the meanings of words, phrases and sentences." (EX-1014, 50 (emphasis in original).)

### B. Interactive/Spoken Dialog Systems

39. At a high level, the task of an interactive dialog system "is to enable and support the spoken interaction between the human user and the service offered

by the application system." (EX-1018, 19.) This high-level tasks leads to a number of sub-tasks performed by the system, including:

> - coherence of the user input has to be verified, taking into account linguistic and task- or domain-related knowledge;
> - communicative and task goals have to be negotiated with the user, and problems occurring during the interaction have to be resolved;
> - references like anaphora or ellipses in the user's utterances have to be resolved;
> - inferences which are reasonable in the communicative and task context have to be drawn, and the most probable user reaction has to be predicted.

(EX-1018, 19.)

40. Möller generally presents four types of systems that can be differentiated "[d]epending on the complexity of the dialogic interaction":

- **Command systems:** "They are characterized by a direct and deterministic interaction. To each stimulus from one agent corresponds a unique response from the other. The response is independent of the state or context of each agent. This type of interaction is normally not considered as a dialogue, and is called a 'tool metaphor'. Example: Pressing a key on the keyboard results in a character appearing on the screen." (EX-1018, 19.)
- **Menu dialog systems:** "To this class belong simple question-answer user interfaces, where dialogue and task models are merged. The interaction is mainly system-directed, permitting only very little user initiative (e.g. barge-

in). In contrast to command systems, several exchanges may be necessary in order to provoke one action of the application system. On the other hand, one user input can provoke different responses, depending on the internal state of the system, e.g. the current level in the menu structure. Example: So-called 'Interactive Voice Response' (IVR) systems which enable an interaction via Dual Tone Multiple Frequency (DTMF) or keyword recognition." (EX-1018, 19.)

- **Spoken Dialog System (SDS):** "This narrow class of systems disposes of distinct and independent models for task, user, system, and dialogue. Context information is taken into account using a particular knowledge base or dialogue history. Multiple types of references can be processed. An SDS may be capable of reasoning, of error or incoherence detection, internal correction, anticipation, and prediction." (EX-1018, 19-20.)
- **Multimodal dialogue systems including speech:** "Systems of this class show the same characteristics as SDSs do, but in addition they are able to process and synchronize different modality information. Appropriate modalities for the individual interactions have to be selected, both at the system input and output side." (EX-1018, 20.)

41. An interactive/spoken dialog system has a "modality" which "is simply a form (or mode) of representing information as output from, or input to, a

computer system (Hovy and Arens, 1990).” (EX-1014, 77.) Some systems only support a single modality (e.g., speech) whereas others support multiple and are referred to as multimodal systems.

42. Spoken dialog systems, including multimodal dialog systems, are “an interface between a human user and an application system which uses speech as the interaction modality (Fraser, 1997).” (EX-1018, 18.) As such, the interface must be able to process information “coming from and going to the user through the speech-technology-based interface (voice user interface, VUI)” and information “coming from and going to the application system through a specialized (e.g., SQL-based) interface.” (EX-1018, 18.) The system “must achieve a number of actions in order to be able to give a response, and the response will depend on the internal state of the system, or on the context of the interaction.” (EX-1018, 18.)

43. Most early spoken dialogue systems were designed for and operated over the telephone network, largely due to the widespread availability and standardized nature of telephony infrastructure. These systems typically employed a phone server interface to manage user access and interactions via voice calls, making them suitable for tasks like automated customer service and information retrieval. (EX-1018, 20.)

44. A spoken dialog system “consists of six major components which are accessed by the user” via an access device. (*See, e.g.*, EX-1018, 20.) These are:

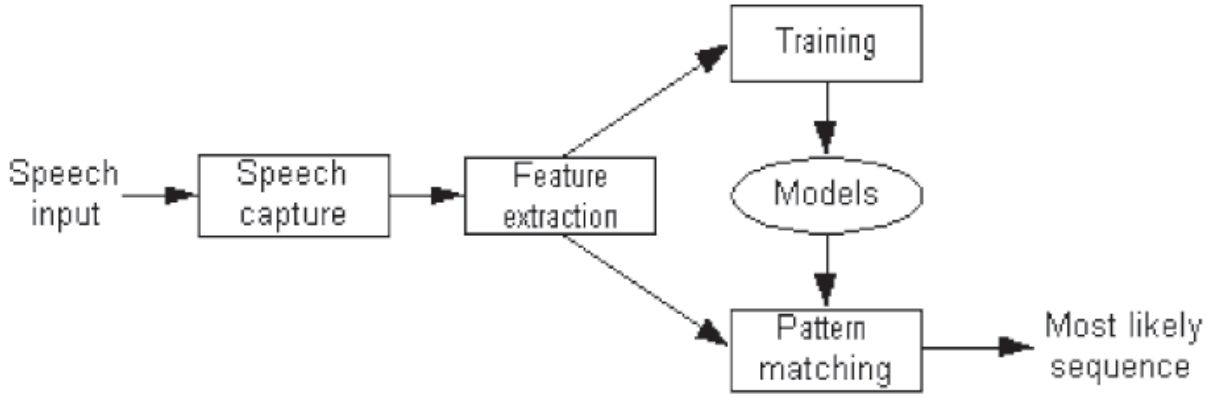
- **Phone Server:** Provides access to the user over the telephone network.
- **Speech Recognizer:** Converts the spoken input into a word string or word hypothesis graph.
- **Semantic Analyzer:** Interprets the recognized word string semantically, producing a semantic frame.
- **Dialogue Manager:** Interprets the semantic frame in context, manages dialogue history, and controls dialogue flow.
- **Response Generator:** creates natural language replies based on the dialogue manager's output.
- **Speech Synthesizer:** Converts the response into speech, and transmits it back to the user.

45. In the next section I describe what are the key components of any speech dialogue system: the speech recognizer, the semantic analyzer, the dialogue manager, and the response generator and speech synthesizer.

### 1. Speech Recognition Component

46. The speech recognition component is commonly referred to as an automatic speech recognizer (ASR). The speech recognition process is illustrated by McTear's Figure 4.2 (below). Speech recognition begins when a user "speaks into a microphone or telephone handset." (EX-1016, 83.) The analog acoustic signal "is captured by the microphone and converted to an electric current, which is then

passed to an analogue-to-digital converter within the sound card that translates the current into a stream of bits that represent the sound." (EX-1016, 83.) This part of the process, known as signal processing, "extract[s] the relevant set of features from the acoustic signal." (EX-1016, 83.)




**McTear, Figure 4.2**

47. The extracted features are then "classified as sequences of phonemes using an acoustic model, and these sequences of phonemes are then combined into words that represent the system's estimate of what the user said." (EX-1016, 83.) The process of phoneme and word identification involves an acoustic model "that shows how each word consists of a sequence of phonemes and how each phoneme relates to the values of the features extracted from the acoustic signal" and a "language model that specifies permissible sequences of words." (EX-1016, 84.)

48. "The acoustic model captures variability in pronunciation using probabilities." (EX-1016, 84.) Words may be pronounced in different ways by

different individuals. (*See, e.g.*, EX-1016, 84.) For example, the word "tomato" may be pronounced "T AH M EY T OW" by American speakers or "T AH M AA T OW" by British speakers. (EX-1016, 84.) Hidden Markov Models ("HMMs") are used in speech recognition systems "to represent [] variable patterns of speech." (EX-1016, 85.)

### 2. Language Understanding

49. The language understanding component "analyse[s] the output of the speech recognition component" and "assign[s] a meaning representation that can be used by the dialogue manager." (EX-1016, 91.) Traditionally, language understanding involved "syntactic analysis, to determine the constituent structure of the recognised string (i.e., how the words group together), and semantic analysis, to determine the meanings of the constituents." (EX-1016, 91.) Language understanding is problematic because of ambiguity in natural language and ill-formed input. (EX-1016, 91.)

50. McTear lists the following ways in which natural language can be ambiguous:

- **Lexical ambiguity***:* "A word may belong to more than one part of speech, for example, 'book' can be a noun or a verb." (EX-1016, 91.) This type of ambiguity "can usually be resolved within the context of the other words in the sentence, as in the sentence 'book a flight to London'." (EX-1016, 91.)

- **Sense ambiguity:** "A word can have different meanings, for example, 'bank' can be a financial institution or the side of a river." (EX-1016, 91.)
- **Structural ambiguity:** "The relationship between the phrases in the sentence is ambiguous, for example, 'a flight to London arriving at 9.' On a purely syntactic analysis, there are two possible readings, one in which the flight arrives at 9, and the other in which London arrives at 9." (EX-1016, 91.)

51. In typical systems, the semantic analyzer uses predefined grammars, slot-filling methods, or statistical models to extract relevant information from user utterances. These might include details such as dates, locations, or actions, which are organized into semantic frames or attribute-value structures. For instance, in a travel booking system, an utterance like "I want to fly to Berlin next Monday" would be converted into a frame containing values for destination and travel date. (*See* EX-1018, 29-30.)

52. The language understanding component will often rely on a component known as a parser to assist in this process. The parser is a tool or algorithm that analyzes the grammatical structure of the recognized speech input. Its role is to break down the sentence into its syntactic components (like subjects, verbs, objects, etc.) and determine how those components relate to each other. This helps the system understand what the user means, not just what words were spoken. For

example, if a user says, “Book a flight to Paris next Friday,” the parser helps identify that:

“Book” is the action (verb),

“a flight” is the object,

“to Paris” indicates the destination,

“next Friday” refers to the date.

This structure can then be processed by the dialogue manager to determine the next system action. (*See* EX-1018, 26, 29-30.)

### 3. Dialog Management Component

53. In a spoken dialog system, the dialog manager is the “central component.” (EX-1016, 107.) The Dialog Manager “accepts spoken input from the user, produces messages to be communicated to the user, interacts with external knowledge sources, and generally controls the dialogue flow.” (EX-1016, 107.) Put differently, according to Möller, “It is the task of the dialogue manager to guarantee the smooth course of the dialogue, so that it is coherent with the task, the domain, the history of the interaction, with general knowledge of the ‘world’ and of conversational competence, and with the user.” (EX-1018, 27.)

54. Dialog managers access both short-term and long-term knowledge to operate effectively. Short-term knowledge, sometimes referred to as short-term memory, often is used to refer to information particularly pertinent to a particular

session whereas long-term knowledge, sometimes referred to as long-term memory, often is used to refer to information that would be relevant across multiple user sessions. Because of this characteristic of these two types of knowledge, it is common for short-term knowledge to be non-persistent and long-term knowledge to be persistent. Dialog managers, like all computational systems, have incorporated the concepts of short-term and long-term knowledge long before the priority date of U.S. Patent No. 8,886,536 ("'536 patent").

#### 4. Response Generator and Speech Synthesizer

55. The response generator and speech synthesizer in a spoken dialogue system is responsible for converting internal representations—typically derived from the dialogue manager—into linguistically and pragmatically appropriate output. "They are described together, because the strict separation into a component which generates a textual version of the output for the user (response generation) and another one which generates an acoustic signal from the text (speech synthesizer) is not always appropriate." (EX-1018, 33.)

56. As Möller explains, "[r]esponse generation involves decisions about what information should be given to the user, how this information should be structured, and about the form of the message (words, syntax). (EX-1018, 33.) The system must decide not just what to say, but how to say it—balancing informativeness, clarity, and conversational appropriateness. This process may be

implemented via formal grammars or simpler "template" approaches, in which sentences are generated by filling predefined structures with contextually relevant content.

57. Möller further notes that at each dialogue act, "the response generator builds a template sentence ... filling gaps from the content of the current semantic frame, the dialogue history, and the result of the database query." (EX-1018, 33.) Higher-level generation rules govern how much information can be included in a single utterance and how it should be structured, particularly when there is too much content to communicate all at once. Importantly, response generation should be informed by the user model, adapting language and content based on the user's background knowledge and experience. As Möller puts it, this module must "respect the user model, e.g. with respect to his/her expected domain knowledge and experience." (EX-1018, 33.) Ultimately, the response generator ensures that the system communicates in a way that is coherent, concise, and aligned with user expectations.

### C. Using Domain Agents to Improve Spoken Dialogue Systems

58. As increasingly advanced computational systems developed in the 1980s and 1990s, researchers sought to move beyond one-size-fits all unitary systems by distributing dialogue tasks among specialized modules or "agents." A

significant early milestone was the Open Agent Architecture ("OAA") developed at SRI International (initially reported in Cohen in 1994). (EX-1019, 1-7).

59. OAA provided a general framework for integrating distinct software agents in a distributed environment. It introduced a facilitator (or, what it called, "Server") that brokered communication between agents via a shared goal blackboard. This allowed multiple agents – for instance, one for speech recognition, one for language understanding, and others for different domain tasks – to collaborate on fulfilling a user's request. The OAA and similar agent-based frameworks in the 1990s demonstrated the feasibility of delegating subtasks to domain-specific agents in a dialogue system.

60. In parallel, the MIT Spoken Language Systems Group was developing the GALAXY architecture. Galaxy implemented a hub and spoke system directly into a spoken dialogue system. This system provided for a central hub routed messages to and from various specialized servers (agents). (EX-1020, 192-95.) Each server was responsible for a particular function or "domain" (for example, parsing, database query, travel information, weather updates, etc.). By 1996, MIT researchers had built dialogue systems in multiple domains (such as air travel planning and weather inquiry) using this architecture. (EX-1020, 192-95.) Then, by 1998, the same group developed the GALAXY-II system, a refined version used as the backbone of the DARPA Communicator program. (*See* EX-1021.) The

Communicator architecture enabled various institutions (including MIT, Carnegie Mellon University, and SRI) to plug in their own domain modules and speech components to a shared hub, making the system truly distributed. The trick to these systems was to develop a common semantic representation scheme (*supra* §III.B.2) that allowed "plug-and-play" adaptability across different specialized agents.

61. Shortly thereafter, in 1999, Bor-shen Lin and Colleagues announced "a distributed architecture for cooperative spoken dialogue agents." (*See* Lin.) In this design, each agent managed the dialogue for a single domain, and a top-level interface agent coordinated them. When a user's input was received, the system would automatically decide which domain agent should handle it, allowing domain switching within one conversation. The following diagram illustrates the architecture of this system:

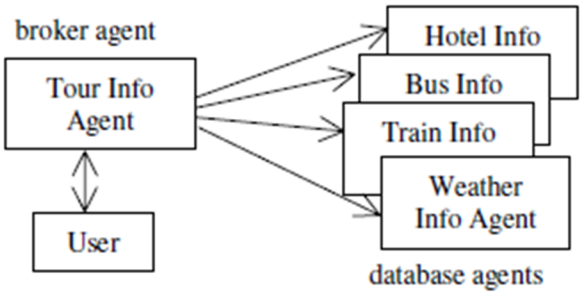



**Figure 1.** An example of centralized model for tour information service

**Lin, Figure 1**

62. By 2005, numerous systems had been developed using the multi-domain-agent architecture in speech dialogue systems. Two of the most prominent

examples are O'Neill in 2004 ("O'Neill (2004)"; EX-1022) and Pakucs in 2003 (EX-1023). Indeed, in a 2010 review article, Nakano and colleagues remarked how "[s]o far many researchers have tackled multi-domain spoken dialogue systems." (EX-1024, 75.)

63. O'Neill (2004)'s system, developed at Queen's University in Belfast, included "cross domain dialogue modelling" which included the use of a discourse history that maintains a record of user-system interaction *across* domains in the form of a stack. (EX-1022, 1,3.) This allowed for consistent dialogues across the multiple domains, a critical advancement for any truly conversational system.

### D. Intent-Based Advertising

64. Intent-based advertising refers to the strategic targeting of individuals who are actively seeking information—often in relation to a specific product or purchase decision. This method of advertising hinges on the ability to discern commercial intent from user behaviors, particularly search activity. (EX-1025, 1.) As Dai describes it, understanding whether a person is "about to commit or in the middle of a commercial activity" can significantly enhance the relevance and profitability of content delivery across the web. (EX-1025, 1.)

65. The concept is not new: traditional mediums like newspaper classifieds and trade directories have long allowed advertisers to place messages directly in front of consumers demonstrating clear intent to engage in commercial activity. The

evolution continued with the widespread adoption of telephone directories, most notably the Yellow Pages. The Yellow Pages became an iconic example of intent-based advertising in the pre-digital era, allowing businesses to list themselves in categories that matched consumer searches precisely. People who needed a plumber, restaurant, or mechanic would actively consult these directories, creating a natural alignment between consumer intent and advertised services.

66. The rise of digital technology led to services like 1-800-FREE-411 in the early 2000s, an innovative phone-based directory assistance service that provided business listings free to callers. Funded by brief audio advertisements, this service tapped directly into consumer intent—callers explicitly seeking information about businesses or services received relevant advertisements. By embedding advertising into directory assistance calls, 1-800-FREE-411 effectively anticipated and served immediate consumer needs. Similar services included GOOG-411, a free voice-activated search tool provided by Google, and Tellme Networks, which also offered voice-driven business search with ad-supported content. These services bridged the gap between analog information-seeking and digital convenience, all while monetizing the moment of user intent.

67. While the idea is not new, technology provides tools to improve the precision with which digital platforms can now detect and respond to these signals of interest. Indeed, as of the earliest priority date of the ’536 family patents,

researches were actively studying how to improve such tools. For example, Dai describes numerous frameworks for determining commercial intent, such as:

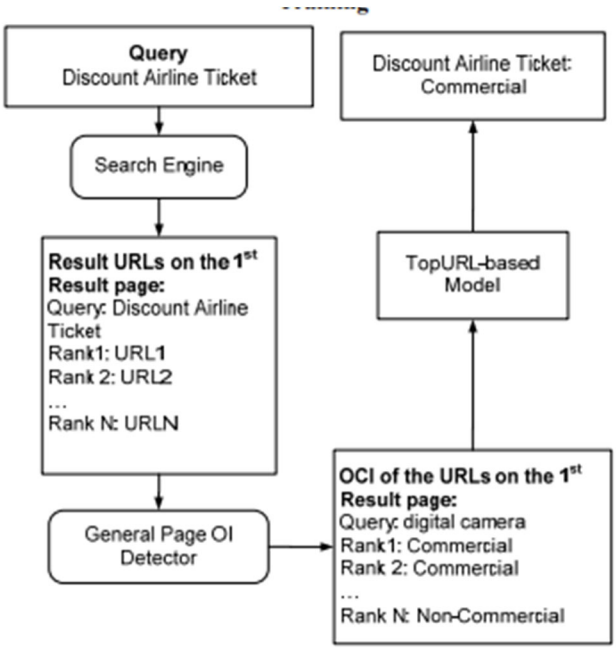



**Figure 5: Detect OCI based on Search Result Landing Pages - Prediction**

**Dai, Figure 5**

## E. Web-Based Advertising Tools

68. Some of the fastest growth of web-based advertising took place between the years of 1995 and 2005, where total online ad revenue increased from approximately $55 million to $12.5 billion, an increase of over 200x. (EX-1026, 11.) As discussed further below, the period between 1995 and 2005 corresponded with the development of new technologies that took advantage of the unique capacity of web-based platforms to provide consumers with contextually appropriate advertisements that were suited for the particular consumer. (EX-1027 7-29.)

69. This technology addressed a long-standing practical issue in advertising. As one investigator observed in 2004, "[e]nsuring that media is bought effectively and not directed towards the 'wrong people' has always been a challenge for marketers." (EX-1028, 1.) Thus, advertisers' "objectives are often to target advertising to specific consumer groups." (EX-1028, 1.)

**1. Early Web-Advertising Systems**

70. Around 1995, early websites and search engines, such as Yahoo! and AltaVista, offered basic forms of advertising that matched ads to the general topic of a page. For instance, a user browsing a page about travel might see banner ads for airline deals or hotel bookings. (*See* EX-1027, 9.) These ads were relatively unsophisticated, relying primarily on static keywords and general page categories rather than user behavior or preferences. They were also often manually placed and lacked precision.

71. A key development in these early systems came with ad platforms like DoubleClick, which started in 1996 (and was bought by Google in 2007). These platforms allowed advertisers to not only deploy ads across many websites but to also systematically track the effectiveness of ads. As Google described in the press release announcing the acquisition: DoubleClick's ad servers placed ads in the "right location at the right time" and reported real-time performance, which allowed advertisers to "change the content, and timing of ads almost on the fly." This made

it possible to "maximize the return on investment for a given ad or campaign." (EX-1029, 4.)

72. Another development from this period was the development and implementation of tools that allowed search-based targeting, sometimes called "sponsored search."

73. Sponsored search is a form of online advertising in which advertisers pay to have their messages displayed alongside search engine results, targeting users based on the keywords they enter. These features enabled search engines to offer their core services for free while generating substantial revenue through paid listings. As Fain and Pedersen explain, sponsored search integrates "relevance-targeted text advertisements" directly into the search experience, and by 2005, search engines were delivering over 13 billion sponsored results per week. (EX-1030, 1.)

74. The model is built on several key components: advertiser-submitted content (titles, descriptions, and URLs) tagged with relevant keywords; a bidding system where advertisers set prices they are willing to pay per click; and an auction-based ranking mechanism that determines ad placement based on both bid value and, in many systems, performance metrics such as click-through rate (CTR). (EX-1030, 2-3.)

75. GoTo.com, later known as Overture Services, was the first to successfully deploy this model in 1998. It allowed advertisers to bid for placement

on specific keywords, and search engines would display these listings in rank order based on the bid amount. (EX-1030, 1-2.) Relevance was ensured through a combination of editorial review and, increasingly, automated algorithms.

76. The model quickly gained traction. Google's adaptation of sponsored search in 2002 introduced a significant innovation: ranking ads not solely by bid amount but by the product of bid and CTR, thereby rewarding ads that were not only lucrative but also engaging to users. (EX-1030, 3.) This reflected one of the earliest ways in which advertising engaged in a feedback loop based on which ads drove customer clicks. Yahoo!, having acquired Overture in 2003, continued to refine the system by introducing the "Click Index," a measure comparing a listing's observed CTR to expected norms. (EX-1030, 3.)

77. Beyond basic cost-per-click (CPC) pricing, the industry expanded to include cost-per-action (CPA) and cost-per-mille (cost per thousand impressions or CPM) models. These enabled advertisers to align their expenditures more closely with measurable outcomes, such as purchases or signups, and allowed publishers to better monetize their platforms. (EX-1030, 2.) With the rise of behavioral targeting, time-of-day bid adjustments, and sophisticated analytics, sponsored search evolved into a dynamic, data-driven marketing platform.

78. One of the most famous instantiations of this approach was Google Adwords, which launched in 2000. (EX-1031, 1.) AdWords enabled advertisers to

"purchase individualized and affordable keyword advertising" that appeared "instantly on the google.com search results page." (EX-1031, 1.) With this system, ads were "highly targeted," triggered only when a user entered a search query that matched an advertiser's "carefully selected keywords." (EX-1031, 1-2.) As Google emphasized, this approach "ensures advertisers will be delivered a prequalified visitor and will not pay for 'off target' impressions." (EX-1031, 2.)

79. The system also offered detailed "web-based reporting tools," including real-time "click-through rates and number of impressions delivered," and "visitor interest bar[s]" to provide immediate feedback on ad effectiveness. (EX-1031, 2.) Ads appeared as "quick-loading AdWords text ads … to the right of the Google search results." (EX-1031, 2.)

**2. User Modelling Approaches**

80. By 2007, basic data analysis tools and growing web tracking methods had been developed that allowed websites to gather information from which advertisers could build a model of who was viewing the webpage. Using these models, advertisers could then substantially improve their ability to show ads that were likely to get a response. As Plummer explains, such targeting methods have been shown to significantly improve key performance indicators such as click-through and conversion rates, offering a clear incentive for advertisers to invest in increasingly granular audience segmentation. (EX-1027, 13.)

81. The inventors of the ’536 patent describe three “models” that can similarly inform the selection of an advertisement: The “personalized cognitive model,” the “generalized cognitive model,” and the “environmental model.” These models track onto three broad categories of modelling used in advertising, including web-based advertising.

82. First, the ’536 patent describes an environmental model that includes “user location, user activity, track user actions, and/or other environmental information…” (EX-1001, 5:27-29).

83. The described “environmental model” closely tracks a range of advertising strategies that deliver messages based on attributes of a user’s immediate physical or temporal context. Two primary examples are geotargeting—the use of location data to determine where a user is at the moment an ad is served—and daypart targeting, which tailors ad delivery according to the time of day. (EX-1027, 17-24.)

84. The ’536 patent also describes a personalized cognitive model that includes “statistical abstracts of user interaction patterns.” (EX-1001, 5:34-36.)

85. This described model closely tracks the idea of behavioral targeting, which refers to the practice of delivering advertising based on inferences drawn from a specific user’s prior online activity. (EX-1027, 11-16.) This activity may include websites visited, search queries submitted, purchases made, or other interactions that

form a behavioral profile. The goal is to increase ad relevance by aligning content with the user's demonstrated interests, rather than their current context or demographic characteristics. (EX-1027, 11-12; *see also* EX-1032.)

86. This method emerged as a key innovation in digital advertising in the early 2000s, as improvements in browser tracking, data storage, and network-level advertising allowed for more sophisticated observation of user behavior. According to Duane Morris LLP (2004), behavioral targeting became feasible through the collection of clickstream data, which enabled advertisers to infer a user's interests without requiring that the user explicitly provide personal data. (EX-1032, 2). This approach allowed networks to serve ads that were customized based on individual behavioral signals, rather than generalized audience traits.

87. Plummer describes behavioral targeting as one of the most powerful tools available to online advertisers, noting that networks can deduce user profiles based on patterns of activity across unrelated websites. For instance, a user who visits several indie music blogs, parenting forums, and auto comparison sites in a short time span may be categorized as having multiple distinct interests—each of which can trigger targeted ad delivery. (EX-1027, 12 (citing industry practice circa early 2000s).) Plummer also cites empirical findings indicating that behaviorally targeted ads can achieve significantly higher engagement than standard or contextual ads. One cited study found that such ads had "108% higher click-through rates than

in-context advertising" and a 19% increase in post-click conversion rates, suggesting not just improved attention but also increased action. (EX-1027, 13 (citing Tacoda and Revenue Science data, ca. 2004–2006).)

88. Plummer further situates behavioral targeting within a broader strategic framework. It explains that behavioral methods can be used both for acquisition (e.g., targeting users in the early stages of purchase consideration) and for retention (e.g., re-targeting users who abandoned a shopping cart). Plummer emphasizes that the strength of behavioral targeting lies in its ability to match ad messaging to users' demonstrated, rather than assumed, interests—effectively bypassing the limitations of self-reported data or static segment definitions.

89. The '536 patent also describes creating models of "relevant characteristics of the user" including "demographics." (EX-1001, 6:41-43.)

90. This idea closely tracks the idea of Demographic Targeting, or the delivery of advertising based on consumer characteristics such as age, gender, household income, education level, or occupation. (EX-1027, 8-9.) These attributes are typically inferred or obtained through user registration data, third-party data aggregators, or predictive modeling. The central premise is that certain demographic groups exhibit consistent preferences or purchasing behaviors, making them suitable proxies for audience selection.

91. Plummer identifies demographic targeting as one of the earliest and most established forms of audience segmentation. (EX-1027, 8.) In the context of online advertising, it represents a digital extension of traditional media planning practices where ad buys were matched to the assumed demographics of television programs, radio shows, or print publications. Online platforms adapted these methods by either collecting self-reported data (e.g., during account creation) or applying statistical models to infer demographic attributes from behavioral or contextual signals. (EX-1027, 8-9.)

### 3. Using Analytics to Improve Models

92. Burns noted in the early 2000s that "[a]ll the search engines track click-through rates (CTRs) and impression data for each of your keywords and creatives." (EX-1033, 251.) These metrics serve as a foundational component of modern advertising systems, capturing both the visibility of advertisements (impressions) and the corresponding user responses (clicks). The systematic collection of this data enables advertisers and platforms to evaluate ad performance and optimize delivery strategies. Beyond campaign reporting, this interaction data provides a valuable input for refining the user models described in the previous section, contributing to ongoing adjustments in how audiences are segmented and targeted.

93. User profiles in digital advertising systems are constructed through the aggregation of interaction data over time. These interactions may include ad clicks,

search queries, content views, and responses to promotional offers. Systems use this data to develop statistical models that estimate a user's interests, preferences, or likelihood to engage with specific content. Plummer explains that advertising networks such as Tacoda and Revenue Science observe behavioral patterns across multiple unrelated sites to infer topic-level interests. These observations allow the system to assign users to affinity segments—clusters defined by similar patterns of interaction—without requiring explicit input from the user. (EX-1027, 12-13 (citing Tacoda and Revenue Science).) The result is a dynamic profile that evolves with continued interaction and is used to inform future ad selection and delivery.

94. This profiling process forms the basis of behavioral targeting systems, which rely on interaction data to infer user preferences and predict future behavior. As described in the previous section, behavioral targeting involves constructing statistical models that associate observed actions—such as ad clicks or content consumption—with underlying interests. These models are not static; rather, they are continuously updated as new interaction data becomes available. Each new data point serves to refine the system's estimates, enabling more accurate classification of user interests and improving the relevance of subsequent ad delivery. This iterative updating mechanism is central to maintaining model accuracy over time and adapting to shifts in user behavior. As Plummer describes, the behavioral approach is not just about making ads more relevant, the approach also allows

companies to track and analyze customers to "profile[] each by their brand interactions." (EX-1027, 53 (citing Anfuso 2005).) The result of this is that the company's "database will only become richer over time as behavioral tracking continues with each successive year." (EX-1027, 54.)

## IV. '536 Patent

### A. Overview of the '536 Patent

95. The '536 patent claims priority to non-provisional U.S. Application No. 11/671,526, filed on February 6, 2007, now U.S. Patent No. 7,818,176. Although I do not have an opinion about whether the '536 patent is entitled to this priority date, each applied reference was filed or published before February 6, 2007.

96. The '536 patent is titled "System and Method for Delivering Targeted Advertisements and Tracking Advertisement Interactions in Voice Recognition Contexts."

97. Figure 3 of the '536 patent, which I reproduce below, is an embodiment of the disclosed method for delivering targeted advertisements and tracking advertisement interactions (EX-1001, 7:27-29):

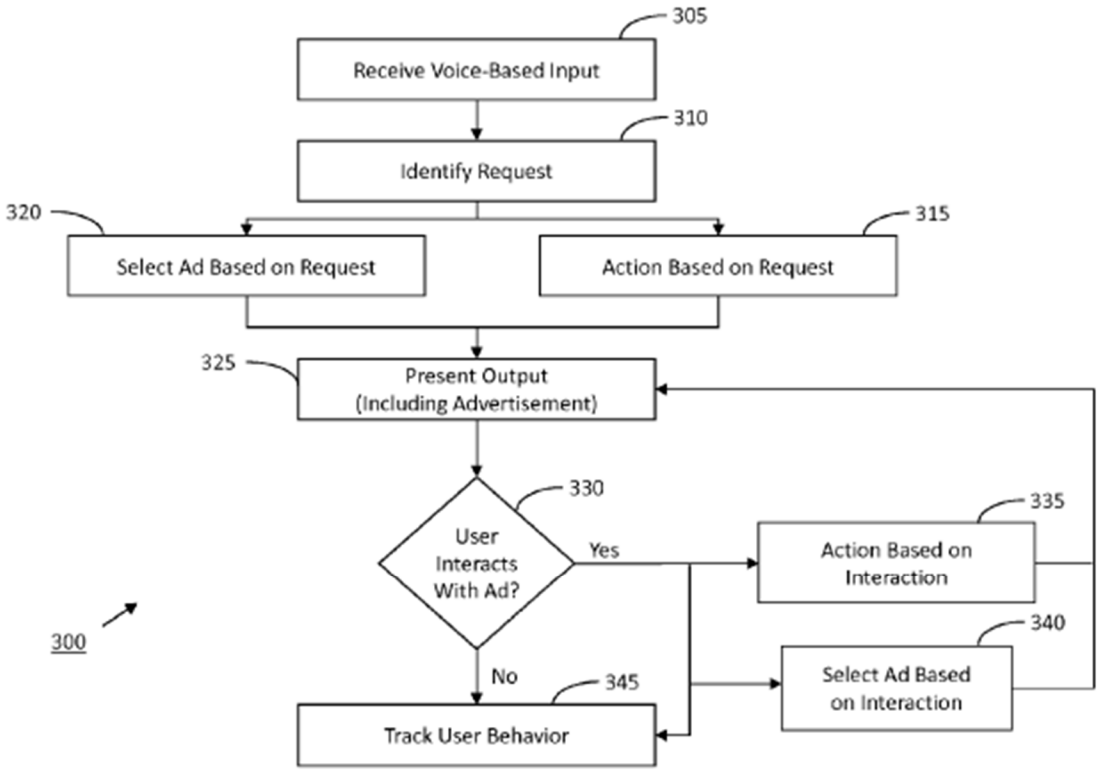



Figure 3

**'536 Patent, Figure 3**

98. Figure 3 of the '536 patent teaches "an operation 305, where a voice-based input, including at least a user utterance" is received at a "voice-user interface" (EX-1001, 7:29-32.) The voice-based input may include "one or more requests," including "requests to retrieve information, perform tasks, explore or gather information, or otherwise interact with a system or device." (EX-1001, 7:38-42.) The utterance may include "cross-device requests," such as calculating a route with a navigation device and making a telephone call using a mobile phone, and "cross-application requests," such as searching for an address and/or phone number using

a voice search engine and calculating a route using a navigation application. (EX-1001, 7:53-59.)

99. Figure 3 of the ’536 patent teaches that operation 310 identifies the request in the voice-based input (EX-1001, 7:38-39.) Operation 310 interprets the request or requests “based upon previous utterances in a current conversation, utterances in previous conversations, context-based information, local and/or global user profiles, or other information.” (EX-1001, 7:62-65.) “Upon identifying the one or more requests,” Figure 3 shows that “[1] action may be taken based on the identified requests in an operation 315, while [2] one or more advertisements may be selected in an operation 320.” (EX-1001, 8:19-22.)

100. Figure 3 of the ’536 patent shows that “[u]pon taking action in operation 315 (e.g., to service the request) and selecting one or more advertisements in operation 320 (e.g., in relation to the request), an output may be presented to the user in operation 325.” (EX-1001, 8:34-37.) The output in operation 325 may include a “voice-based or otherwise audible response” or it may be displayed on a display mechanism associated with the device. (EX-1001, 10:57-62.)

101. The output in operation 325 “may include requested information, an indication of whether a requested task was successfully completed, whether additional information is needed to service the request (e.g., including a prompt for the information), or other information relating to an action based on the request.”

(EX-1001, 8:38-43.) "[T]he output may also include advertisements, as selected in operation 320." (EX-1001, 8:44-45.)

102. The system and/or method selects advertisements in operation 320 based on various criteria. (EX-1001, 8:52-9:3.) The advertisement may be selected based on "words or other content of the request," "requested tasks/information," "a topic or category associated with the requested tasks/information," "location information," or "user specific and/or global user profiles." (EX-1001, 8:52-9:3.) "[A]dvertisers may pay a premium to prioritize an advertisement in relation to similar advertisements," and "statistical profiles" may be used to "define affinities between advertisements, topics, users, etc. (e.g., based on click-through or conversation rates, or other tracking information)." (EX-1001, 10:44-50.)

103. The system and method of the '536 patent may also select advertisements to "advance the conversation towards the goal." (EX-1001, 9:35-37.) In one example of this, the '536 patent teaches that operation 325 may locate nearby points of interest based on a user's preferred topics when the user requests a navigation route. (EX-1001, 9:37-55.) In another example, where the request includes "incomplete, ambiguous, or unrecognized information, (e.g., 'Do you know [mumbled words] Seattle?')" the '536 patent teaches that the advertisements "may be selected based on user and/or global profiles reflecting an affinity between certain advertisements associated with Seattle and user preferences, profiles, etc." (EX-

1001, 10:6-24.) "[B]y selecting an advertisement, indicating dissatisfaction with an advertisement, or otherwise interacting with an advertisement, the interaction may be used to build context and shared knowledge for a subsequent course of the conversation." (EX-1001, 10:24-28.)

104. The bottom portion of Figure 3 of the '536 patent shows that "[s]ubsequent interaction between the user and the presented advertisements may be monitored in a decisional operation 330." (EX-1001, 11:12-14.) Further tracking takes place in operation 345. (EX-1001, 11:38-39.) "For example, operation 345 may track historical data about users, conversations, topics, contexts, or other criteria to associate information with the selected advertisement." (EX-1001, 11:39-42.)

105. Figure 3 of the '536 patent further shows that "when the user elects to interact with the advertisement, action may be taken based on the interaction in an operation 335." (EX-1001, 11:14-16.) The user's interaction with the advertisement may take various forms, including clicking on the advertisement displayed or requesting action from a voice-enabled media device. (EX-1001, 11:16-31.) "[A]dditional advertisements may be selected in an operation 340 based on the interaction, using similar techniques as described in connection with operation 320 …." (EX-1001, 11:32-34.)

106. There are two other figures in the ’536 patent. Figure 1 of the ’536 patent, which I reproduce below, shows an “exemplary system 100 for implementing a voice user interface.” (EX-1001, 3:18-20.)

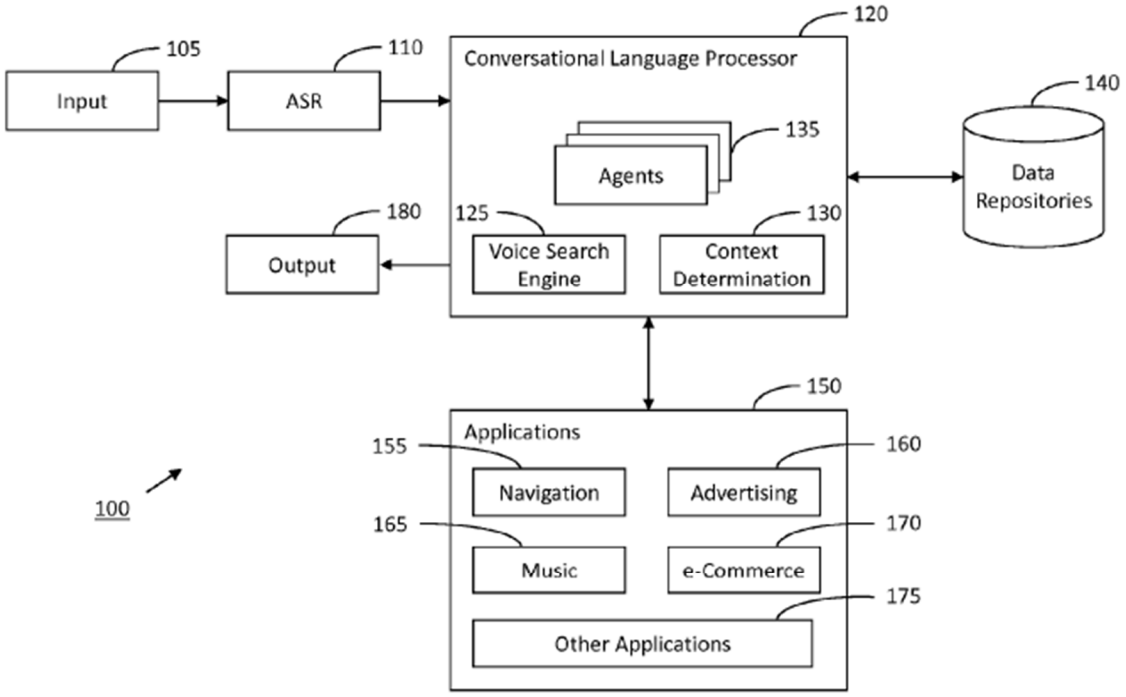



Figure 1

**‘536 Patent, Figure 1**

107. The ’536 patent teaches that System 100 may be used to “control navigation devices, media devices, personal computers, personal digital assistants, or any other device supporting voice-based inputs.” (EX-1001, 3:22-24.)

108. System 100 includes an input device 105 that receives a user input, “including at least a voice-based user utterance.” (EX-1001, 3:52-54.) Input devices

may include a microphone or other multi-modal input, such as a touch-screen interface, keypad, or other input. (EX-1001, 3:56-59.)

109. System 100 includes a speech recognition engine 110, which the 536 patent refers to as an "Automatic Speech Recognizer" or "ASR." The ASR processes the utterance received at the input device 105 and generates one or more "preliminary interpretations" of the utterance using "various techniques," including "phonetic dictation to recognize a stream of phonemes," (EX-1001, 3:59-65), or "dictionary and phrase tables," (EX-1001, 3:40-43).

110. The ASR's preliminary interpretations are provided to a conversational language processor 120. (EX-1001, 4:10-11.) As shown in Figure 1 of the '536 patent, the conversational language processor 120 includes "a voice search engine 125, a context determination module 130, and one or more agents 135, among other things …." (EX-1001, 4:10-14.)

111. The context determination module 130 "establish[es] a context for the current utterance by having domain agents 135 competitively generate a context-based interpretation of the utterance (e.g., by scoring possible interpretations and selecting a highest scoring interpretation)." (EX-1001, 4:50-55.) The domain agents 135 model various domains, including navigation, music, a specific user, global users, advertising, and e-commerce. (EX-1001, 4:55-59.)

112. The conversational language processor 120 is also capable of "adaptive misrecognition," meaning that it may reinterpret a current utterance or one or more previous utterances. (EX-1001, 5:4-6.)

113. The conversational language processor 120 may be "communicatively coupled" to "one or more data repositories 140 and one or more applications 150." (EX-1001, 4:16-18.) The applications 150 may include a navigation application 155, an advertising application 160, a music application 165, an electronic commerce application 170, and/or other applications 175. (EX-1001, 4:31-34.) The data repositories 140 and the applications 150 "may reside locally (e.g., on a user device), remotely (e.g., on a server), and/or hybrid local/remote processing models may be used (e.g., lightweight applications may be processed locally while computationally intensive applications may be processed remotely)." (EX-1001, 4:34-41.)

114. The conversational language processor 120 generates a "domain-specific conversational response" that is "returned to the user as an output." (EX-1001, 4:18-20.) It may also build "long-term and/or short-term shared knowledge in one or more knowledge sources." (EX-1001, 4:42-44.) "[K]nowledge sources may include information about previous utterances, requests, and other user interactions to inform generating an appropriate response to a current utterance." (EX-1001, 4:44-47.) "The shared knowledge may include public/non-private (i.e.,

environmental) knowledge, as well as personal/private (i.e., historical) knowledge." (EX-1001, 4:47-49.)

115. Figure 2 of the '536 patent, which I reproduce below, shows an "exemplary advertising system 200." (EX-1001, 5:38-439.)

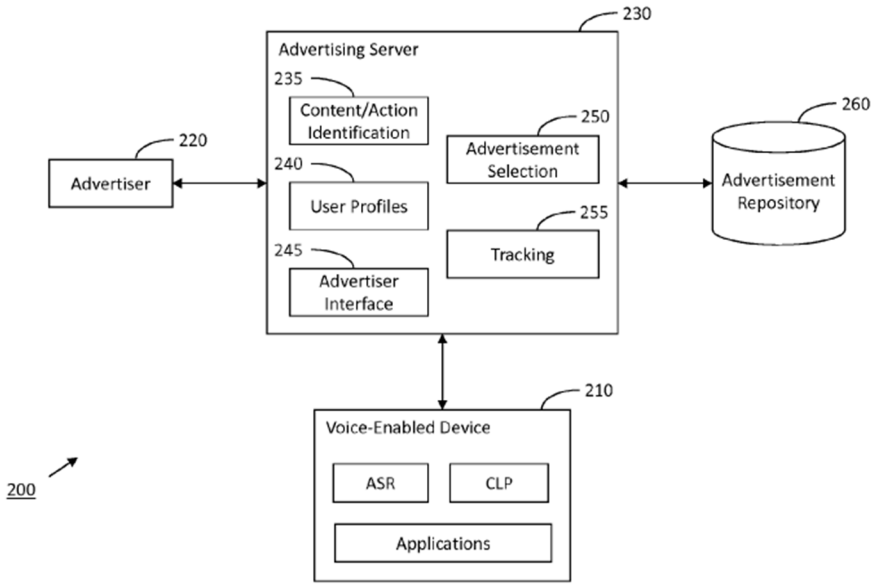



Figure 2

**'536 Patent, Figure 2**

116. System 200 includes an advertising server 230 "for receiving one or more advertisements from an advertiser 220, wherein the advertisements may be stored in a data repository 260 associated with the server 230." (EX-1001, 5:40-43.)

117. System 200 includes a voice-enabled device 210 that establishes communications with the advertising server 230. (EX-1001, 5:48-50.) "Subsequently, advertising server 230 may select one or more advertisements from

among the advertisements stored in data repository 260, and the selected advertisements may be provided to the voice-enabled device for presentation to a user." (EX-1001, 5:50-54.)

118. System 200 teaches that advertisers 220 will access the advertising server 230 via an advertiser interface 245. (EX-1001, 55-56.) "Advertisers 220 may upload targeted advertisements to server 230 via advertiser interface 245, and server 230 may store the advertisements in data repository 260." (EX-1001, 5:56-59.) "[A]dvertisements may include sponsored messages or marketing communications, calls to action, purchase opportunities, trial downloads, coupons, or any other suitable marketing, advertising, campaign, or other information …." (EX-1001, 5:43-48.) The advertisements may be "graphically-based advertisements that include banners, images, audio, video, or any suitable combination thereof." (EX-1001, 5:59-61.)

119. The advertising server 230 of System 200 may include different components, including "a content/action identification module 235, a user profile module 240, an advertisement selection module 250, and a tracking module 255." (EX-1001, 6:4-7.)

120. The content/action identification module 235 of the advertising server 230 identifies "content of the voice-based input (e.g., words in the input), requested information (e.g., search results, a web page, music, video, graphics, or other

information), requested actions (e.g., calculating a navigation route, placing a telephone call, playing a song, etc.), a category or topic related to the input (e.g., music, business, stocks, sports, navigation, movies, etc.) or other criteria to use in selecting an advertisement." (EX-1001, 6:17-25.)

121. The user profile module 240 of the advertising server 230 identifies "characteristics of a specific user (e.g., demographics, personal preferences, location-based information, etc.), global user profiles (e.g., demographic profiles, click-through rates, etc.), or other criteria to use in selecting an advertisement." (EX-1001, 6:25-29.)

122. The advertisement selection module 250 of the advertising server 230 "extract[s] relevant information from the voice-based input" and "select[s] one or more advertisements relevant to the voice-based input based on information extracted using content/action identification module 235 and/or user profile module 240." ('526 patent, 6:10-16.) The advertising server 230 communicates the selected advertisements to the voice-enabled device 210. (EX-1001, 6:44-51.)

123. The tracking module 255 of the advertising server 230 tracks the user's subsequent interactions with an advertisement. (EX-1001, 6:66-7:3.) "[T]racking module 255 may determine whether a conversion or click-through occurs for each advertisement presented to users." (EX-1001, 7:1-3.) "[T]he tracking information may be used to bill or invoice advertisers 220, as well as to improve subsequent

performance and relevance of advertisements selected using advertisement selection module 250." (EX-1001, 7:20-24.)

### B. Level of Ordinary Skill in the Art

124. I understand that certain issues relating to the validity of the '536 patent should be judged from the perspective of a person of ordinary skill in the relevant art, as I discuss below. I have been asked to define the level of a "person of ordinary skill in the art" or "POSITA" at the time the alleged invention as claimed was made. In deciding the level of ordinary skill, I have considered the following factors, which I have been informed are relevant to the determination of the level of ordinary skill in the art.

- levels of education and experience of persons working in the field;
- types of problems encountered in the art;
- prior art solutions to these problems;
- rapidity with which innovations are made; and
- sophistication of the technology.

125. In my opinion, a person of ordinary skill in the art 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.

## C. Prosecution and Litigation History

### 1. U.S. Prosecution Before the U.S. Patent & Trademark Office

126. The ’536 patent issued from U.S. Patent Application Serial No. 14/016,757, filed September 3, 2013.

127. The originally filed application contained 20 claims. (EX-1002, 24-28.) During prosecution, the applicants cancelled those claims in a preliminary amendment and added claims 21-55. (EX-1002, 45-54.)

128. The Examiner rejected originally filed claims 21-25, 30-35, 39-46, and 55 as anticipated by U.S. Patent Application Publication No. 2007/0203736 (“Ashton”). (EX-1002, 65.)

129. Applicants filed a Response on May 9, 2014, cancelling claims 24-26 and amending the remaining claims to incorporate the subject matter that the Examiner had stated was allowable in the first office action. (EX-1002, 85-109.) For example, pending independent claim 21 was amended to add the limitations shown below, where the added limitations are underlined and the deleted limitations are struck through:

21. (**Currently Amended**) A computer-implemented method of ~~obtaining conversationally-related~~ 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:

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

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

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

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

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

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:

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;

obtaining a first interpretation of the second natural language utterance from the first domain agent;

obtaining a second interpretation of the second natural language utterance from the second domain agent; and

determining the interpretation based on one or more of the first interpretation or the second interpretation;

determining, by the one or more physical processors, promotional content based on the interpretation; ~~the first natural language utterance and the second natural language utterance;~~ and

presenting, by the one or more physical processors, the promotional content to a user.

(EX-1002, 86-87.)

Applicants argued that the amendments mooted the § 112 rejections for failure to include an antecedent basis. (EX-1002, 110.)

130. The Examiner allowed the pending claims (21-23 and 27-78) on June 5, 2014. (EX-1002, 295-300.) After the Notice of Allowance, Applicants filed an amendment to correct "minor errors" in independent claims 56, 59, 77, and 78, changing the phrase "a user" to "the user." (EX-1002, 363.) The Examiner issued a Supplemental Notice of Allowability on September 8, 2014. (EX-1002, 383-84.)

### 2. Inter Partes Review before the Patent Trial and Appeal Board

131. The '536 patent was subject to two IPR petitions. (EX-1002, 440, 469.)

132. Claims 1, 2, 7-13, 16-23, 25, 27, 30, 39, 41, 42, 44, 47, 48, and 50-52 of the '536 patent were challenged in a first petition on three grounds:

| Claims Challenged | Statutory Basis | References |
|---|---|---|
| 1, 13, 16, 25, 27, 30, 39, 44, 50 | 103(a) | U.S. Patent Application Publication 2004/0193420 ("Kennewick")<br><br>Japanese Patent No. 2002-297626A ("Yonebayashi") |
| 2, 7–12, 17, 18, 21–23, 41, 42, 47, 48, 51, 52 | 103(a) | Kennewick, Yonebayashi, and U.S. Patent No. 6,859,776 ("Cohen") |
| 19, 20 | 103(a) | Kennewick, Yonebayashi, Cohen, and U.S. Patent No. 6,526,335 ("Treyz") |

133. For the first ground, petitioner relied primarily on Kennewick and referenced Yonebayashi as "additional teaching regarding providing promotional content based on the interpretation of natural language utterances and presenting promotional content to a user." (EX-1002, 456.) The PTAB concluded that the petitioner failed to demonstrate a reasonable likelihood of prevailing on the first ground (Kennewick + Yonebayashi) because the petitioner failed to show that Kennewick discloses: (1) domain agents that interpret the detected speech, as required by claims 1, 13, 16, 25, 27, and 30; (2) multiple domain agents, as required by claims 1, 13, 16, 25, 27, and 30; (3) identifying multiple requests associated with a single utterance, as required by claims 39 and 44; and (4) obtaining a user's profile information that specifies prior interaction with "items," including advertisements, as required by claim 50. (EX-1002, 456-65.) PTAB found that the proposed combination of Kennewick and Yonebayashi failed to disclose determining promotional content based on the first and second utterances, as required by claim 50. (EX-1002, 465-66.)

134. For the second ground, petitioner relied on Kennewick and Yonebayashi, as set forth in the first ground, and relied on Cohen to teach "performing an activity based [on] a received utterance, recognizing reference words, and initiating a purchase transaction." (EX-1002, 466.) PTAB found that petition failed to establish that Kennewick, Yonebayashi, and Cohen satisfied the

claim limitations for the same reasons as set forth in deciding the first ground. (EX-1002, 466.)

135. For the third ground, petitioner relied on Kennewick, Yonebayashi, and Cohen, as set forth in the first and second grounds, and relied on Treyz to teach "calculating a navigation route, determining promotional content based on a navigation route, and selecting a media item and different promotional content based on the media item." (EX-1002, 466-67.) PTAB found that petition failed to establish that Kennewick, Yonebayashi, Cohen, and Treyz satisfied the claim limitations for the same reasons as set forth in deciding the first ground. (EX-1002, 467.)

136. Claims 1, 3-6, 13-16, 24-40, 43-46, 49, 50, and 53-55 were challenged in a second petition on one ground:

| Claims Challenged | Statutory Basis | References |
|---|---|---|
| 1, 3–6, 13–16, 24–40, 43–46, 49, 50, 53–55 | 103(a) | Kennewick, Yonebayashi |

137. The PTAB concluded that petitioner demonstrated a reasonable likelihood of prevailing in showing that claims 33-36 would have been obvious over the combination of Kennewick and Yonebayashi. (EX-1002, 489-493.) In making this determination, PTAB concluded that Kennewick disclosed a system that "receives three utterances from a user and how a third such utterance is used to reinterpret a second utterance," and a system that "can provide location-based

marketing programs." (EX-1002, 492.) PTAB further concluded that "Yonebayashi discloses a system that provides tailored advertisements to the user and enables interaction between the user and the advertisement" and "allows the user to interact with the system to provide supplemental or substitute promotional content per a user's request." (EX-1002, 492-93.)

138. However, the PTAB concluded that petitioner failed to establish a reasonable likelihood of prevailing in showing the unpatentability of 31 of the 35 challenged claims. (EX-1002, 504.) The PTAB grounded its decision in the same rationale as the set forth in the first institution decision, namely that the combination of Kennewick and Yonebayashi failed to disclose domain agents that interpret detected speech and multiple domain agents, as required by claims 1, 3-6, 13-16, and 24-31 of the '536 patent; failed to disclose multiple requests associated with a single utterance, as required by claims 39, 40, 43, 44-46, and 49; and failed to disclose obtaining a user's profile information that specifies prior user interaction with items, as required by claims 50 and 53-55. (EX-1002, 483-89, 498-500.) PTAB also concluded that Kennewick and Yonebayashi did not disclose reinterpreting an initial ambiguous utterance based on a user's subsequent interaction with an advertisement, as required by claims 37 and 38. (EX-1002, 496.)

139. Therefore, the PTAB exercised its discretion under 35 U.S.C. § 314(a) and declined to institute IPR on the second petition. (EX-1002, 504.)

### 3. Proceedings in the U.S. District Court

140. Claims 1-55 of the '536 patent were asserted against Amazon in a patent infringement action in the U.S. District Court for the District of Delaware in Case No. 1:19-cv-01410-MN.

141. During those proceedings, certain terms in the '536 patent were subject to claim construction, where the following constructions were agreed upon or otherwise imposed by the Court:

| Claim Term | Joint Proposed Construction |
|---|---|
| "domain agent"<br><br>'536 patent, claims 1, 4, 32 | "software with domain-specific behavior and information" (EX-1010, 1-2) |
| "speech recognition engine" / "speech recognition"<br><br>'536 patent, claims 1, 10, 32 | "software or hardware that recognizes the words or phrases in the natural language utterance" (EX-1010, 2) |
| "context"<br><br>'536 patent, claim 3 | No construction necessary (EX-1010, 2 (rejecting Amazon's proposal that the term should mean "subject domain(s) that applies to a user utterance")) |

## V. GROUND 1: THE COMBINATION OF O'NEILL AND JONES RENDERS OBVIOUS CLAIMS 1-12, 16-23, 30, 32 AND 50-55

### A. Overview of the Combination

142. The combination of U.S. Pub. No. 2005/0278180 to O'Neill, et al. ("O'Neill"; EX-1005) and U.S. Pub. No. 2007/0174258 to Jones, et al. ("Jones"; EX-1006) discloses every limitation of claims 1-12, 16-23, 30, 32 and 50-55. O'Neill

and Jones are analogous art to the '536 patent because both references are directed to speech recognition systems and software.

143. I describe both O'Neill and Jones further below.

**1. O'Neill**

144. O'Neil discloses "[a] system for conducting a dialogue with a user." (EX-1005, Abstract.) O'Neill provides the following high-level overview of the disclosed process:

> [A] method of conducting a dialogue with a user comprising the steps of:
>
> (a) receiving a communication from the user;
>
> (b) 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;
>
> (c) storing information from the communication in one or more data-storage units;
>
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
>
> (e) forwarding a response to the user; and
>
> (f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.

(EX-1005, [0052]-[0058].)

145. O'Neill describes the "individual processing units are adapted to perform tasks of different specificities" and "processing units with more specialized

processing rules are substantially independent of those with less specialized processing rules." (EX-1005, [0029]-[0032].)

146. The system employs a cohort of domain agents which it labels as "Experts" or "service agents":

> The domain-specific functionality of the improved dialogue manager is provided by a cohort of agents. 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 or sub-dialogue and elicits information that is stored in a specialized dialogue frame. However, by using inheritance a common approach to dialogue management is established, wherein each agent inherits the same confirmation strategy independent of its domain specialty. In contrast with the simple agent designs of prior art dialogue managers, the cohort of agents employed in the improved dialogue manager are specifically designed to collaborate with each other to detect and facilitate user-led changes in conversational topic.

(EX-1005, [0073].)

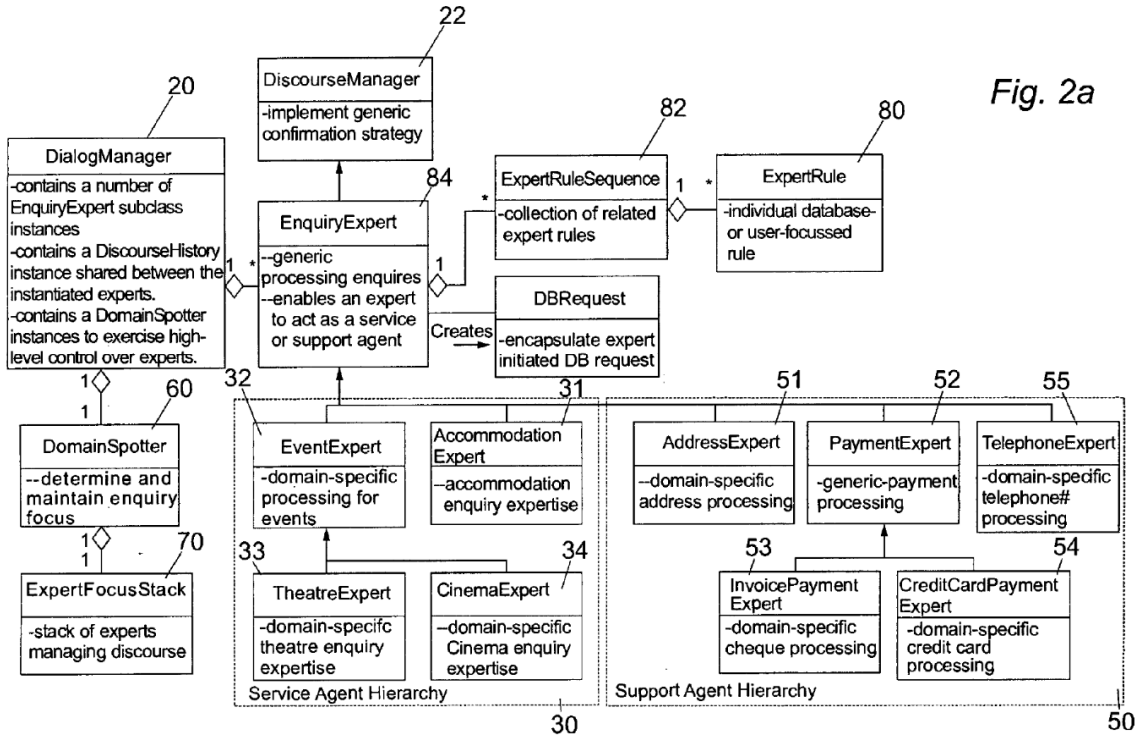
147. For example, there are agents for accommodation booking, event booking, and payment processing:

> In the present example, the class of service agents 30 include an AccommodationExpert 31 and an EventExpert 32 that are respectively responsible for searching for and booking accommodation and event tickets. In addition, the class of support agents 50 include an AddressExpert 51 and a PaymentExpert 52 that are respectively responsible for acquiring the user's address and payment details. It 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, [0123].)

148. An exemplary schematic of the O'Neill system in provided in Figure 2a:




**O'Neill, Figure 2(a)**

149. As depicted in the figure, the system includes a DomainSpotter which processes natural language utterances from the user to determine which domain agent is most appropriate to handle the user's enquiry.(EX-1005, [0198].) O'Neill provides the following explanation of what is depicted in Figure 2a:

> In particular, referring to FIGS. 2*a* and 2*b*, the discussion will consider the following topics:
>
> > (1) the identification of the agent most suited with dealing with a particular dialogue topic at the start of, and during a dialogue session;
> >
> > (2) the implementation of the generic confirmation strategy by the Discoursemanager and its inheritance by domain-specific Experts;
> >
> > (3) the creation new informational elements in response to confirmed data provided by the user;
> >
> > (4) the transfer of dialogue control from a service agent to a support agent;
> >
> > (5) the detection and transition rules for the implementation of user-initiated transfers of dialogue control;
> >
> > (6) the operation of a failsafe mechanism for identifying a handling agent; and
> >
> > (7) the implementation of the expert rules encapsulated within the domain-specific Experts.

(EX-1005, [0188]-[0195].)

150. O'Neill discloses that "[t]he first step in starting a dialogue session is to identify the most appropriate agent to process a user's enquiry. The agent-identification process is typically performed automatically by a DomainSpotter 60." (EX-1005, [0198].)

151. O'Neill discloses the details of how the DomainSpotter operates:

> Referring to FIGS. 2*a* and 2*b*, the improved dialogue manager 20 uses a DomainSpotter 60 to identify an

> appropriate handling agent for a user's query. In an initial attempt to identify a handling agent the DomainSpotter 60 focuses its attention on service agents 30 only and supplies each service agent 30 with the output of the semantic parse received from the natural language parser.
>
> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60.
>
> The service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific heuristics to the more detailed processing of the user's enquiry. For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast.

(EX-1005, [0200]-[0202].)

152. O'Neill provides the following example of how the DomainSpotter operates in practice:

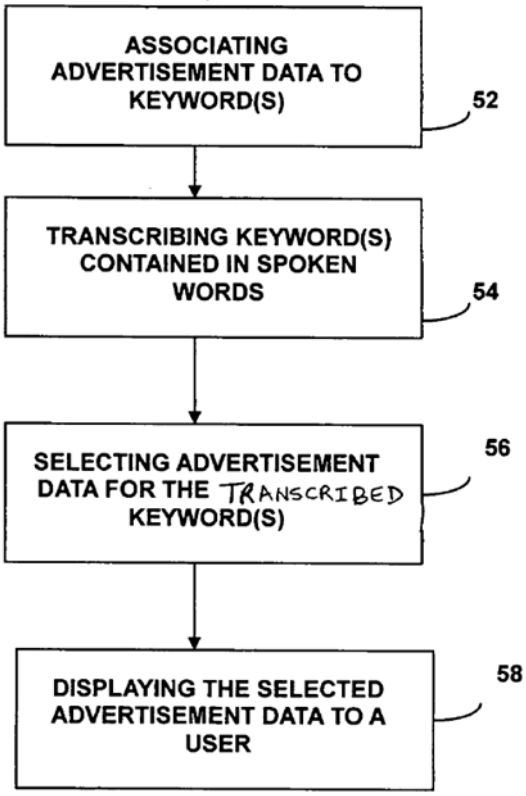
> In addition, specialized agents also attain a higher score for specialized parser tags than generic agents. For example, a user request such as "I'd like to go to see THE LION KING 2." might parse as event_enquiry: [Eventtype][Movies].THE LION KING 2.
>
> In this case, although the EventExpert 32 could award a score for the event_enquiry, the CinemaExpert 34, as a child of EventExpert 32, would award a score not only for the event enquiry, but for Movies as well, and so would be the winner.

(EX-1005, [0203]-[0204].)

153. O'Neill also provides a helpful example of a prolonged dialogue in an interactive session between the system and a user that is exemplary as to how the system operates in practice that, for purposes of brevity, I do not recreate here in its entirety, but I do refer to below for specific disclosures. (EX-1005, [0287]-[0422].)

**2. Jones**

154. "The system and method disclosed [in 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].) Specifically, "[t]argeted advertisement is provided to mobile device users" wherein "users may submit voice requests," "relevant keyword(s) are determined from the requests and corresponding advertisements and/or products associated with the keyword(s) are provided to the users while the users are waiting for responses to the requests." (EX-1006, Abstract.) Such features are depicted in Figure 2 shown below.




**Jones, Figure 2**

155. The system uses information associated with the user to determine what advertisement to provide, such as the user's phone number or the user's GPS location. (*See* EX-1006, [0031].) "When a voice query from one of the user devices 12 is received by the system 10, for example, for a search without requesting assistance from a provider (or guide), the query is transcribed and passed to one or more search engines or resources, and advertisements associated with the query may be selected based on keyword histograms in the database 18 that "best fit" the query and as appropriate for the user telephone number. For example, when a user requests information pertaining to 'restaurants in Indiana' and data records in the database 18

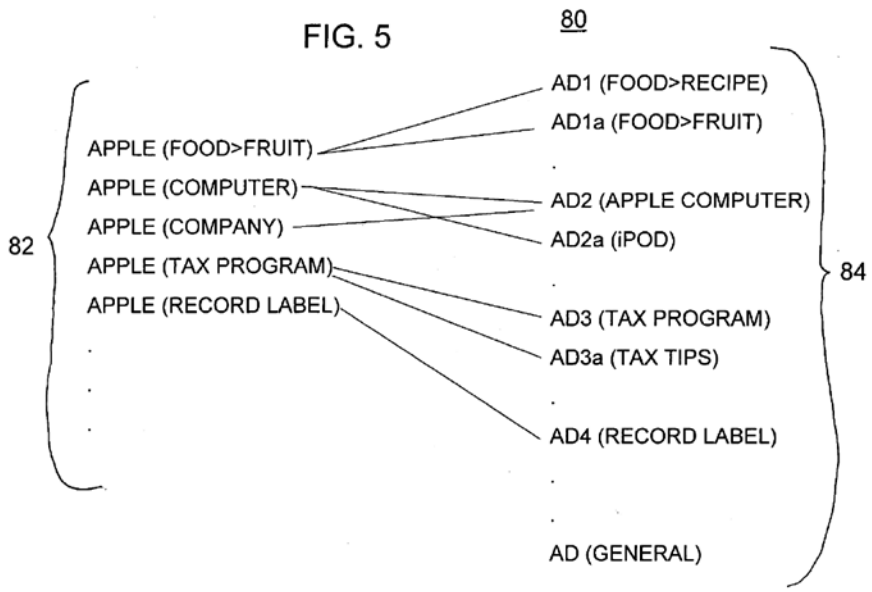
indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other geographic data such as GPS coordinates may also be used to select from restaurant advertisements in Indiana to provide one in or near the designated geographic location." (EX-1006, [0031].)

156. To determine the appropriate advertisement to provide to users, the system has a database that maintains information about the user's prior requests and what information was provided in response to those prior requests: "The database 18 may maintain information about … requests submitted from the user devices 12, results generated by the providers in response to the requests, advertisements presented to the users while corresponding requests are being processed by the system 10 including advertisements previously presented for requests pertaining to similar subject matter." (EX-1006, [0032].) The system also tracks whether users interact with certain advertisements after making requests with certain keywords: "For example, the database 18 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].)

157. The database also maintains information regarding user engagement with advertisements that is used by the system when selecting future advertisements: "For example, the server 16 may compute 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 (this information may be delivered back to the server 16 from the advertiser), etc." (EX-1006, [0033].) The system may also select advertisements based on information about the user that is associated with the phone number and/or the user, i.e., a user profile. (EX-1006, [0063].)

158. The database also may maintain a histogram of usages for keywords or phrases with multiple meanings and/or forms. (EX-1006, [0053].) The system may prompt the user for clarification or "the system 10 may classify or categorize keywords and/or phrases based on information from external sources such as Internet directories (e.g., DMOZ), online dictionaries and/or encyclopedias (e.g., Webopedia, Wikipedia, etc.) or other similar sources." (EX-1006, [0053].) Figure 5 shows the association of qualified keywords with advertisements:




**Jones, Figure 5**

159. However, the system may also use the context of the user's search itself to determine the meaning of a phrase that may have ambiguity: "For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes." (EX-1006, [0040].)

### 3. Motivation to Combine

160. While O'Neill discloses the limitations of asserted claims related to receiving natural language utterances from a user, processing the natural language utterances using specific processing rules using domain agents and responding to the natural language utterances, O'Neill does not describe using the information learned

from the user during the conversation to provide the user with targeted advertisements or promotional content. Jones provides those details regarding determining and providing the user with promotional content related to the natural language utterances.

161. 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.

162. A POSITA would have been motivated to combine O'Neill and Jones for a number of reasons.

163. First, Jones suggests the combination: Jones states: "The system 10 may perform speech to text conversion using either speech transcription using human transcribers or using conventional speech-to-text processing, also known as automatic speech recognition (ASR)." (EX-1006, [0025].) O'Neill also suggests the system can be combined with other systems: "The improved dialogue manager will be described by way of example in an accommodation and event booking and payment system. It will be appreciated 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 or other transactional systems." (EX-1005, [0089].)

164. Second, the systems disclosed in O'Neill and Jones are designed to help users with requests. Jones discloses that the purpose of the system is to "receive[] voice requests or queries submitted by users or information seekers from user devices 12, deliver[] the requests to providers for processing and return[] responses resulting from the processing by the providers to the users." (EX-1006, [0023].) Similarly, O'Neill discloses a system that "receives a communication from a user, assigns the communication to a processing unit according to the communication's semantic content and stores information relating to the communication in a data storage unit; and the assigned processing unit processes the communication in accordance with the processing unit's processing rules and provides a response to the user." (EX-1005, Abstract; *see also* EX-1005, [0052]-[0058].) Because the systems are both designed to accomplish the same goal—i.e., process and respond to user requests—a POSITA would be motivated to combine their teachings. Moreover, both Jones and O'Neill are systems that are designed to assist users with similar types of requests. As discussed above in §§V.A.1-2, Jones and O'Neill both assist users with requests related to identifying and connecting with businesses, which would further motivate a POSITA to combine their respective teachings to enhance the combined system.

165. Third, the addition of the targeted advertising of Jones to the interactive dialogue system disclosed by O'Neill would provide opportunities to monetize

and/or increase revenues of the interactive dialogue system, which the system in O'Neill is already designed to handle. The system in O'Neill is designed to have an interactive dialogue with the user and assist the user with certain requests, such as booking a hotel room or book a cinema ticket. (*See, e.g.*, EX-1005, [0132]-[0154]; [0264].) To facilitate the system performing those tasks, O'Neill includes payment processing functionality that the user can interact with to book a hotel room or buy a cinema ticket. (EX-1005, [0366]-[0389].) Therefore, a POSITA would be motivated to combine the teachings of Jones to further monetize the platform disclosed in O'Neill.

166. Fourth, the combination is nothing more than use of a known technique of targeted advertising (Jones) to improve a similar system that utilizes automatic speech recognition (O'Neill's interactive dialogue system). Both Jones and O'Neill operate in the same way, as both systems rely on automatic speech recognition and processing to interact with a user who is requesting information from the systems.

167. The results of the combination of O'Neill and Jones would have been predictable to a POSITA and the POSITA would have had a reasonable expectation of success for at least two reasons.

168. First, both O'Neill and Jones are directed to systems that interact with users who are requesting information by using automatic voice recognition and processing techniques.

169. Second, Jones's method of determining and presenting targeted advertisements to users based on keywords in the user's natural language utterances would be an obvious addition to the O'Neill interactive dialogue system that is already designed to connect users with businesses for commercial purposes (e.g., booking hotels or cinema tickets). It would have been well within the skill set of a POSITA to integrate the targeted advertising system into the interactive dialogue system taught by O'Neill.

### B. Independent Claims 1 and 32

170. Independent claims 1 and 32 include substantially overlapping processing limitations ([1A]-[1L]/[32A]-[32L]), as shown in the table below. The only substantive difference between the claims is that claim 1 claims "*computer-implemented method of providing promotional content related to one or more natural language utterances and/or responses*" and claim 32 claims a "*system for providing promotional content related to one or more natural language utterances and/or responses*." Because of the overlap between the limitations in these claims, I address these limitations together. In this table, I also provide a citation to the section discussing each limitation:

| Claim 1 | Claim 27 | Section |
|---|---|---|
| [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: | [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: | §V.B.1 |
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [32A] receive a first natural language utterance; | §V.B.2 |
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [32B] provide a response to the first natural language utterance; | §V.B.3 |
| [1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance; | [32C] receive a second natural language utterance relating to the first natural language utterance; | §V.B.4 |
| [1D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance; | [32D] perform speech recognition to recognize one or more words of the second natural language utterance; | §V.B.5 |
| [1E] determining, by the one or more physical processors, domain information for the one | [32E] determine domain information for the one or more | §V.B.6 |

| Claim 1 | Claim 27 | Section |
|---|---|---|
| or more recognized words based on the first natural language utterance; | 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: | [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: | §V.B.7 |
| [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; | [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; | §V.B.8 |
| [1H] obtaining a first interpretation of the second natural language utterance from the first domain agent; | [32H] obtaining a first interpretation of the second natural language utterance from the first domain agent; | §V.B.9 |
| [1I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and | [32I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and | §V.B.10 |
| [1J] determining the interpretation based on one or more of the first interpretation or the second interpretation; | [32J] determining the interpretation based on one or more of the first interpretation or the second interpretation; | §V.B.11 |
| [1K] determining, by the one or more physical processors, | [32K] determine promotional content based on the interpretation; and | §V.B.12 |

| Claim 1 | Claim 27 | Section |
|---|---|---|
| promotional content based on the interpretation; and | | |
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [32L] present the promotional content to a user. | §V.B.1 |

### 1. Preamble [1P] / [32P]

[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:

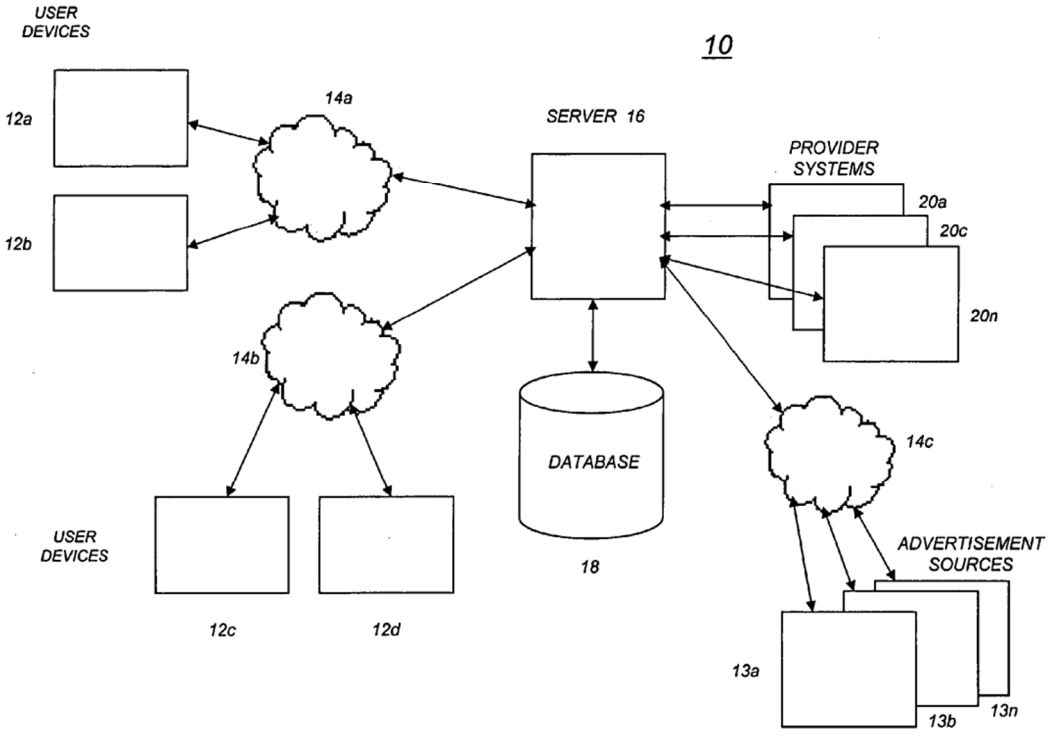
[32P] A system for providing promotional content related to one or more natural language utterances and/or responses, the system comprising: one or more physical processors programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors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

171. To the extent that the preamble is limiting, the combination of O'Neill and Jones discloses "*[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*," and "*[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*" as discussed below in §§V.B.2-12.

172. Jones discloses "*[a] computer-implemented method*," and is "*implemented by a computer system that includes one or more physical processors executing one or more computer program instructions*" and discloses "*the system comprising: one or more physical processors programmed to execute one or more computer program instructions*" as the disclosed system is performed on a computer system that is depicted in Figure 1:

FIG. 1



**Jones, Figure 1**

173. Jones describes the computer system depicted in Figure 1:

> As shown in FIG. 1, system 10 receives voice requests or queries submitted by users or information seekers from user devices 12, delivers the requests to providers for processing and returns responses resulting from the processing by the providers to the users. The user devices 12 shown in FIG. 1 may be mobile phones, landline phones, specialized voice query terminals, voice-enabled PDA's (personal digital assistants), or any other source that allows a user to enter a query via speech entry or a Short Message Service (SMS) and communicate with server 16 via networks 14*a* and/or 14*b*. Although networks 14*a* and 14*b* are illustrated as connecting user devices 12 to the server 16, the system disclosed is not limited to any particular number of networks via which user devices may connect to the server 16. The user devices 12 may also be typical touch-tone telephones, two-way radios, wifi phones, cordless phones, portable media devices, or any communication device that allows the user to communicate with the server 16 via networks 14*a* and/or 14*b*. The communication system can include packet switched facilities, such as the Internet, circuit switched facilities, such as the public switched telephone network, radio based facilities, such as a wireless network, etc.

(EX-1006, [0023].) A POSITA would understand that Jones discloses a computer implemented system.

174. O'Neill similarly discloses a computer implemented system as the system using "processing units" to perform the disclosed method:

> According to the invention there is provided a system for conducting a dialogue with a user comprising:
>
> > an assignment unit, a plurality of processing units each of which comprise one or more processing rules, and a plurality of data storage units;
> >
> > wherein the assignment unit receives a communication from a user, assigns the communication to a processing unit according to the communication's semantic content and stores information relating to the communication in a data storage unit; and
> >
> > the assigned processing unit processes the communication in accordance with the processing unit's processing rules and provides a response to the user.

(EX-1005, [0023]-[0027].)

175. Jones further discloses a "*method of providing promotional content related to one or more natural language utterances and/or responses*": "The system and method disclosed [in 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].)

176. The '536 patent does not define the term "promotional content." In fact, the specification for the '536 patent does not even use the term "promotional content" at all besides in the claims. Therefore, the specification does not provide any instruction on how to define "promotional content." However, the term

"promotional content" would be well understood by a POSITA to include the "advertising" disclosed in Jones.

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

| |
|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance;<br><br>[32A] receive a first natural language utterance; |

177. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a first natural language utterance,*" and "*receive a first natural language utterance*."

178. Specifically, O'Neill discloses:

> a method of conducting a dialogue with a user comprising the steps of:
>
> (a) **receiving a communication from the user;**
>
> (b) 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;
>
> (c) storing information from the communication in one or more data-storage units;
>
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
>
> (e) forwarding a response to the user; and
>
> (f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.

(EX-1005, [0052]-[0058].) As discussed above, O'Neill discloses that the system is performed by "processing units" that are "*physical processors*."

179. O'Neill also discloses explicitly that one of the processing units receives the utterances from the user:

> According to the invention there is provided a system for conducting a dialogue with a user comprising:
>
> an assignment unit, a plurality of processing
>
> units each of which comprise one or more processing rules, and a plurality of data storage units;
>
> **wherein the assignment unit receives a communication from a user**, assigns the communication to a processing unit according to the communication's semantic content and stores information relating to the communication in a data storage unit; and
>
> the assigned processing unit processes the communication in accordance with the processing unit's processing rules and provides a response to the user.

(EX-1005, [0023]-[0027].)

180. The "communications" that are processed by the improved dialogue manager disclosed in O'Neill are "natural language utterances": "**The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances** or as a consequence of the agents' own rules. It is this

process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user." (EX-1005, [0077].)

181. Similarly, Jones discloses receiving and processing natural language utterances, including using "*physical processors*": "The users may **submit voice requests** using mobile phones or other mobile networked devices, relevant keyword(s) are determined from the requests and corresponding advertisements and/or products associated with the keyword(s) are provided to the users while the users are waiting for responses to the requests." (EX-1006, Abstract.)

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

| |
|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance;<br><br>[32B] provide a response to the first natural language utterance; |

182. The combination of O'Neill and Jones discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*," and "*provide a response to the first natural language utterance*."

183. As described above, O'Neill provides a response to the user's first natural language utterance using a physical processor:

> According to the invention there is provided a system for conducting a dialogue with a user comprising:
>
> > an assignment unit, a plurality of processing
> >
> > units each of which comprise one or more processing rules, and a plurality of data storage units;
> >
> > wherein the assignment unit receives a communication from a user, assigns the communication to a processing unit according to the communication's semantic content and stores information relating to the communication in a data storage unit; and
> >
> > **the assigned processing unit processes the communication in accordance with the processing unit's processing rules and provides a response to the user**.

(EX-1005, [0023]-[0027].) O'Neill further discloses:

> a method of conducting a dialogue with a user comprising the steps of:
>
> > (a) receiving a communication from the user;
> >
> > (b) 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;
> >
> > (c) storing information from the communication in one or more data-storage units;
> >
> > **(d) using one or more of the processing rules of the assigned processing unit to process the communication;**
> >
> > **(e) forwarding a response to the user;** and
> >
> > (f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.

(EX-1005, [0052]-[0058].)

184. O'Neill provides additional details regarding the system's ability to provide responses to users using physical processors. O'Neill discloses the use of what it labels the "DiscourseManager." O'Neill discloses that "The DiscourseManager 22 is responsible for the improved dialogue manager's 20 overall 'conversational style', (i.e. the generic discourse behavior of the improved dialogue manager 20)." (EX-1005, [0110].) "The DiscourseManager 22 determines the response of the improved dialogue manager 20 to new, modified or negated information from the user." (EX-1005, [0111].)

185. Jones also discloses a system that can, in response to the user's request, "produce automated results." (EX-1006, [0046].)

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

| |
|---|
| [1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance;<br><br>[32C] receive a second natural language utterance relating to the first natural language utterance; |

186. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*," and "*receive a second natural language utterance relating to the first natural language utterance*."

187. O'Neill describes that multiple turns of dialog with the user can be repeated beyond the processing and response to the first utterance, each turn of dialog is received at the "*physical processor*." Specifically, O'Neill states that the dialogue system includes:

> a method of conducting a dialogue with a user comprising the steps of:
>
> (a) receiving a communication from the user;
>
> (b) 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;
>
> (c) storing information from the communication in one or more data-storage units;
>
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
>
> (e) forwarding a response to the user; and
>
> **(f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.**

(EX-1005, [0052]-[0058].)

188. Therefore, a POSITA would understand that the system in O'Neill is designed to receive a "*second natural language utterance relating to the first natural language utterance*" as the "repeating the above steps" in the process would necessarily include additional natural language utterances from the user in the same "dialogue" that are related to the first natural language utterance.

189. O'Neill provides specific examples of dialogue between the user and the system that demonstrates that the system discloses receiving "*a second natural language utterance relating to the first natural language utterance*":

> System: Welcome to the Communicator. How can I help you?
>
> User: I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December.
>
> …
>
> System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Please wait while I check the database ...
>
> …
>
> User: Could you make that a three-star?
>
> …
>
> System: OK, got you. That's three-star accommodation. Is that OK?
>
> …
>
> User: Yes, that's right.
>
> System: Please wait while I check the database.

(EX-1005, [0291]-[0321].)

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

| |
|---|
| [1D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance;<br><br>[32D] perform speech recognition to recognize one or more words of the second natural language utterance; |

190. The combination of O'Neill and Jones discloses "*performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance*," and "*perform speech recognition to recognize one or more words of the second natural language utterance*."

191. I understand that in the Amazon Litigation the parties agreed to construe the term "speech recognition" to mean "software or hardware that recognizes the words or phrases in the natural language utterance." I further understand that the Court adopted that construction. For purposes of my analysis here, I adopt that construction and apply it in my analysis below.

192. As described above, O'Neill describes that multiple turns of dialog with the user can be repeated beyond the processing and response to the first utterance. The system performs "*speech recognition to recognize one or more words*" of each user utterance, including the "*second natural language utterance*." Specifically, O'Neill discloses the following regarding the dialogue system:

> a method of conducting a dialogue with a user comprising the steps of:
>
> (a) receiving a communication from the user;
>
> (b) 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;

(c) storing information from the communication in one or more data-storage units;

**(d) using one or more of the processing rules of the assigned processing unit to process the communication**;

(e) forwarding a response to the user; and

**(f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.**

(EX-1005, [0052]-[0058].)

193. The improved dialogue manager disclosed in O'Neill receives a semantic parse from a natural language parser: "The improved dialogue manager accepts as input semantically tagged key-words and phrases from the Phoenix natural language parser (W. Ward, Proceedings of ICSLP 94, pp. 83-86, Yokohama, 1994). The improved dialogue manager is interfaced to the Galaxy hub 14 through a Java DialogServer class which includes the relevant hub package (importgalaxy.server.*;) and creates and populates frames of information in the Galaxy format." (EX-1005, [0278].) A POSITA would understand that the "Phoenix natural language parser" is used by the system to perform "*speech recognition to recognize one or more words*" using physical processors of all natural language utterances from the user.

194. O'Neill discloses that "a DARPA communicator system is used in an implementation of the present invention, the following section will refer to accompanying FIG. 1 to briefly describe the architecture of this automatic dialogue system." (EX-1005, [0005], [0096].) O'Neill further describes Figure 1 as "an end-to-end automatic dialogue system 1 typically comprises an automatic speech recognizer 2, a natural language semantic parser 4, a decision unit (known as the dialogue manager) 6, a database 'back-end' 8, a natural language generator 10 and a text-to-speech engine 12." (EX-1005, [0013].) Therefore, a POSITA would understand the O'Neill system also includes these components, including an automatic speech recognizer, a natural language semantic parser and a dialogue manager. (*See* EX-1005, [0013].)

195. Additionally, as shown in the example above, the system in O'Neill processes and responds to multiple user utterances, which requires the system to perform speech recognition on each of the utterances in order to respond. (*See* EX-1005, [0291]-[0321].)

196. 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 Based on First Natural Language Utterance [1E] / [32E]

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

197. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, domain information for the one or more recognized words based on the first natural language utterance*," and "*determine domain information for the one or more recognized words based on the first natural language utterance*."

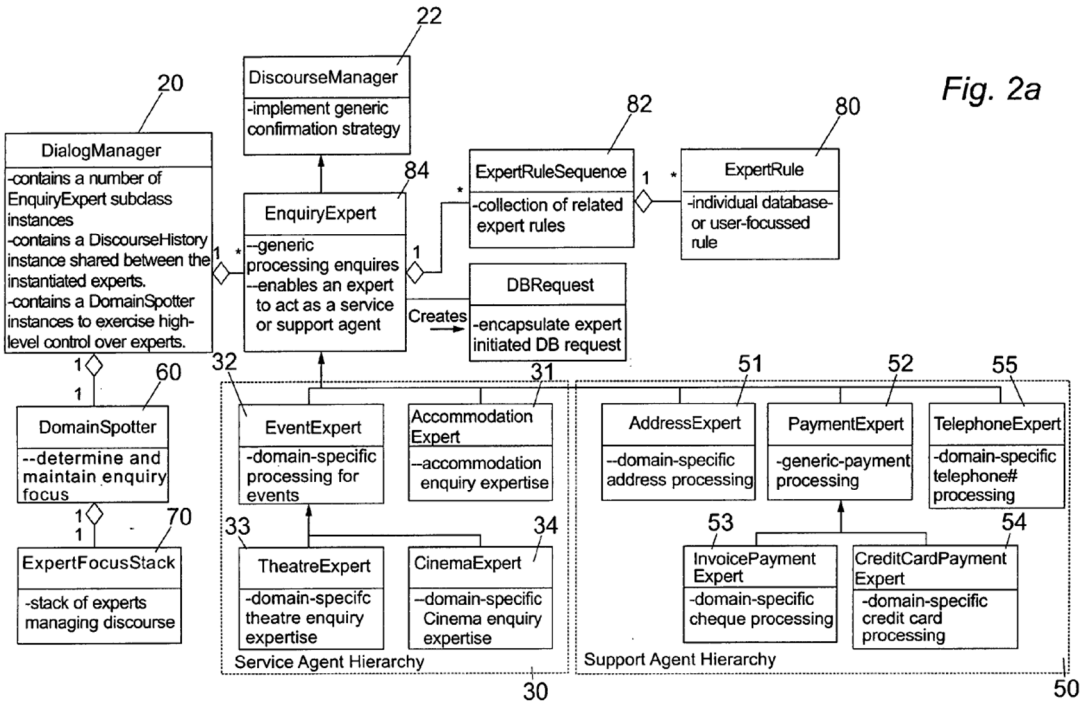
198. O'Neill describes that "[t]he first step in starting a dialogue session is to identify the most appropriate agent to process a user's enquiry. The agent-identification process is typically performed automatically by a DomainSpotter 60." (EX-1005, [0198].)

199. To perform the "*determine domain information*" step, the DomainSpotter in O'Neill uses "*one or more recognized words based on the first natural language utterance*": "In an initial attempt to identify a handling agent the DomainSpotter 60 focuses its attention on service agents 30 only and supplies each service agent 30 **with the output of the semantic parse received from the natural language parser**." (EX-1005, [0200]-[0202].) As discussed above, the system in

O'Neill is performed on a computer system and the DomainSpotter operates using a "*one or more physical processors*." (§V.B.1.)

200. O'Neill describes the "natural language parser" as a prior art parser that parses natural language utterances by semantically tagging key-words and phrases: "The improved dialogue manager accepts as input semantically tagged key-words and phrases from the Phoenix natural language parser (W. Ward, Proceedings of ICSLP 94, pp. 83-86, Yokohama, 1994)." (EX-1005, [0278].) As discussed above, O'Neill discloses that the system also contains an automatic speech recognizer. (EX-1005, [0013].)

201. O'Neill further describes the use of the DomainSpotter: "Referring to FIGS. 2a and 2b, the improved dialogue manager 20 uses a DomainSpotter 60 to identify an appropriate handling agent for a user's query. In an initial attempt to identify a handling agent the DomainSpotter 60 focuses its attention on service agents 30 only and supplies each service agent 30 with the output of the semantic parse received from the natural language parser." (EX-1005, [0200].) This structure is shown in Figure 2a:




**O'Neill, Figure 2a**

202. O'Neill also provides examples of how the DomainSpotter operates: "Referring to FIGS. 2a and 2b, key words from the user's utterance like four-star and hotel fall within the accommodation domain, and are provided with a high score by the AccommodationExpert 31." (EX-1005, [0294].) This example demonstrates that O'Neill determines from a user's first natural language utterance regarding a hotel that the "*domain information for the one or more recognized words based on the first natural language utterance*" is a hotel accommodation.

203. Figure 4 of O'Neill further discloses how a user's utterance is handled by the system after an appropriate handling agent for initiating the dialogue has been identified:

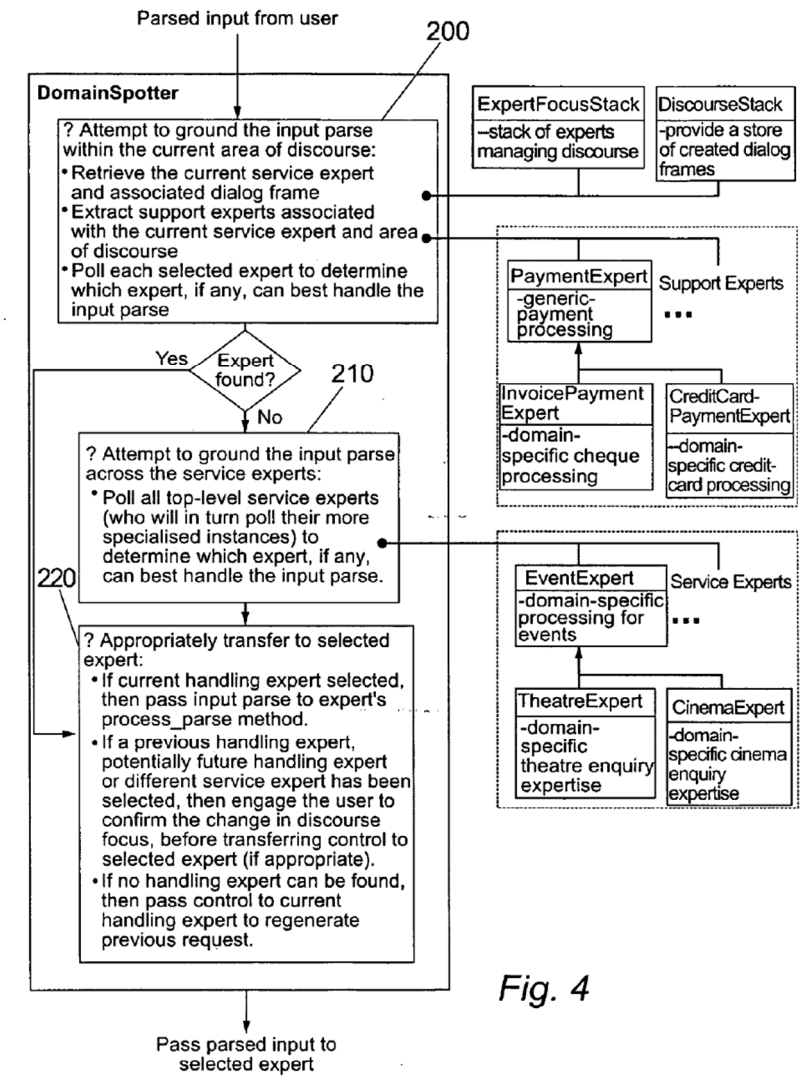



*Fig. 4*

**O'Neill, Figure 4**

204. O'Neill describes Figure 4: "Having identified an appropriate handling agent for initiating a dialogue, FIG. 4 shows the strategy employed by the improved dialogue manager for determining which agent should handle the user's further utterances. **In a first filtering step 200, the improved dialogue manager attempts to restrict the most recent user utterance to the current area of discourse**. In a second filtering step 210, the Dialogue Manager attempts to restrict the user's

utterance to a service enquiry (rather than a support activity). In the final filtering step 220 the Dialogue Manager identifies the appropriate Expert for the user's last utterance." (EX-1005, [0208].) O'Neill is describing using its Dialogue Manager to take the most recent user utterance (which could be, but is not limited to, a "*first natural language utterance*") to identify "the appropriate Expert" which is "*domain information for the one or more recognized words*."

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

[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;

[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:

205. The combination of O'Neill and Jones discloses "*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, wherein processing the one or more recognized words comprises*," and a "*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*," as discussed below in §§V.B.8-10. (*See* EX-1005, [0052]-[0058] ("assigning the communication to one of plurality of processing units, each which comprises one or more processing rules, in accordance with the semantic contention of the communication").)

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

| |
|---|
| [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;<br><br>[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; |

206. The combination of O'Neill and Jones discloses "*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*," and "*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*."

207. I understand that in the Amazon Litigation the parties agreed to construe the term "domain agent" to mean "software with domain-specific behavior and information." I further understand that the Court adopted that construction. For purposes of my analysis here, I adopt that construction and apply it in my analysis below.

208. O'Neil discloses the use of the DomainSpotter to "*provid[e] 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*," which O'Neill refers to as "service agents":

> Referring to FIGS. 2 a and 2 b, the improved dialogue manager 20 uses a DomainSpotter 60 to identify an appropriate handling agent for a user's query. In an initial attempt to identify a handling agent the DomainSpotter 60 focuses its attention on service agents 30 only and **supplies each service agent 30 with the output of the semantic parse received from the natural language parser**.
>
> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, **each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60**.

(EX-1005, [0200]-[0201].) When the DomainSpotter "supplies each service agent with the output of the semantic parse received from the natural language parser" the DomainSpotter is providing the recognized words from the natural language parser to each of the service agents, which are "domain agents" as the service agents are "software with domain-specific behavior and information." For example, O'Neill discloses "[r]eferring to FIGS. 2a and 2b, key words from the user's utterance like four-star and hotel fall within the accommodation domain, and are provided with a high score by the AccommodationExpert 31." (EX-1005, [0294].)

209. O'Neill describes the "domain specific behavior and information" of the "service agents":

> The domain-specific functionality of the improved dialogue manager is provided by a cohort of agents. **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 or sub-dialogue and elicits information that is stored in a specialized dialogue frame**. However, by using inheritance a common approach to dialogue management is established, wherein each agent inherits the same confirmation strategy independent of its domain specialty. In contrast with the simple agent designs of prior art dialogue managers, the cohort of agents employed in the improved dialogue manager are specifically designed to collaborate with each other to detect and facilitate user-led changes in conversational topic.

(EX-1005, [0073].)

210. O'Neill provides examples of these "domain agents" that it refers to as "agents" and "experts":

> In the present example, the class of service agents 30 include an AccommodationExpert 31 and an EventExpert 32 that are respectively responsible for searching for and booking accommodation and event tickets. In addition, the class of support agents 50 include an AddressExpert 51 and a PaymentExpert 52 that are respectively responsible for acquiring the user's address and payment details. It 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, [0123].)

211. O'Neill discloses that the "one or more recognized words" that are provided to the domain agents can be from the "second natural language utterance" as O'Neill discloses that the system performs the following on the "user's last utterance" and on the "user's further utterances":

> Having identified an appropriate handling agent for initiating a dialogue, FIG. 4 shows the strategy employed by the improved dialogue manager for **determining which agent should handle the user's further utterances**. In a first filtering step 200, the improved dialogue manager attempts to restrict the most recent user utterance to the current area of discourse. In a second filtering step 210, the Dialogue Manager attempts to restrict the user's utterance to a service enquiry (rather than a support activity). **In the final filtering step 220 the Dialogue Manager identifies the appropriate Expert for the user's last utterance**.

(O'Neill, [0208].) A POSITA would understand that O'Neill is disclosing that the system provides the "user's further utterances" and "user's last utterance" to the domain agents pursuant to the process described above to determine "which agent should handle the user's further utterances" and identify "the appropriate Expert for the user's last utterance," including a "*second natural language utterance*."

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

| |
|---|
| [1H] obtaining a first interpretation of the second natural language utterance from the first domain agent;<br><br>[32H] obtaining a first interpretation of the second natural language utterance from the first domain agent;<br><br>[1I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and;<br><br>[32I] obtaining a second interpretation of the second natural language utterance from the second domain agent; and |

212. The combination of O'Neill and Jones discloses "*obtaining a first interpretation of the second natural language utterance from the first domain agent*" and "*obtaining a second interpretation of the second natural language utterance from the second domain agent*."

213. O'Neil discloses the use of the DomainSpotter to provide the natural language utterance to the domain agents and "*obtaining a first interpretation of the second natural language utterance from the first domain agent*," and "*obtaining a second interpretation of the second natural language utterance from the second domain agent*":

> Referring to FIGS. 2 a and 2 b, the improved dialogue manager 20 uses a DomainSpotter 60 to identify an appropriate handling agent for a user's query. In an initial attempt to identify a handling agent the DomainSpotter 60 focuses its attention on service agents 30 only and **supplies each service agent 30 with the output of the**

> **semantic parse received from the natural language parser**.
>
> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60.
>
> The service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific heuristics to the more detailed processing of the user's enquiry. **For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast.**

(EX-1005, [0200]-[0202].)

214. A POSITA would understand that when "each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60," each of the service agents (i.e., the domain agents) is interpreting the natural language utterance to determine if it is the appropriate domain agent to handle the user's request. O'Neill also discloses that each domain-specific expert "is able to determine which values have been confirmed (e.g. the user has not challenged an implicit confirmation request by the system) and which have been modified or negate. (EX-1005, [0215].) A POSITA would understand that a determination of a confirmation of 'Yes' in response to the

system's observation 'So, you're staying in Dublin'" is an interpretation of a second natural language utterance. (EX-1005, [0213].)

215. The DomainSpotter applies this process to multiple natural language utterances, including the second natural language utterance. (EX-1005, [0213].) As discussed in the example above, if the user had been asking about hotels in Belfast, the system would eventually determine that the AccomodationExpert is the appropriate domain agent to handle the user's request: "For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast." (EX-1005, [0202].)

216. O'Neill discloses that the "one or more recognized words" that are interpreted by the domain agents can be from the "*second natural language utterance*." O'Neill discloses that the system performs the following on the "user's last utterance" which could be the "*second natural language utterance*":

> Having identified an appropriate handling agent for initiating a dialogue, FIG. 4 shows the strategy employed by the improved dialogue manager for **determining which agent should handle the user's further utterances**. In a first filtering step 200, the improved dialogue manager attempts to restrict the most recent user utterance to the current area of discourse. In a second filtering step 210, the Dialogue Manager attempts to restrict the user's utterance to a service enquiry (rather than a support activity). **In the final filtering step 220 the Dialogue Manager identifies the appropriate Expert for the user's last utterance**.

(EX-1005, [0208].) A POSITA would understand that O'Neill is disclosing that the system performs provides the "user's last utterance" to the domain agents pursuant to the process described above to identify "the appropriate Expert for the user's last utterance," including a "*second natural language utterance*." (*See also* EX-1005, [0261]-[0267] (transferring between domain agents), [0271] (polling domain agents to interpret request).)

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

| |
|---|
| [1J] determining the interpretation based on one or more of the first interpretation or the second interpretation; |
| [32J] determining the interpretation based on one or more of the first interpretation or the second interpretation; |

217. The combination of O'Neill and Jones discloses "*determining the interpretation based on one or more of the first interpretation or the second interpretation*."

218. As discussed above, O'Neill discloses that the domain agent that attains the highest score in degree of relevance to the user's utterance is assigned to handle the user's enquiry, which a POSITA would understand means that the system is "*determining the interpretation based*" on the interpretations of domain agent with the highest score, and then the system assigns that domain agent to handle the user's enquiry. (*See* EX-1005, [0200]-[0202].)

219. O'Neill describes that the domain agents then engage in the dialogue with the user:

> Since the dialogue frame taken from the DiscourseStack 123 details the information previously provided by the user together with the status of the information (repeated, modified, negated, etc.) and the system's previous intentions for confirming or repairing supplied or missing information, the generic confirmation strategy enables the improved dialogue manager 20 to interpret the user's most recent utterance in light of the improved dialogue manager's own last utterance. For instance, if the user answers "Yes" in response to the system's observation "So, you're staying in Dublin", then Dublin can be regarded as the confirmed location.
>
> As will be recalled, the domain-specific 'Experts' all inherit the generic dialogue handling strategies of the DiscourseManager 22. In particular, in order for each Expert to update the record of the evolving dialogue, each Expert takes on the improved dialogue manager's 120 DiscourseHistory 122 as an inheritable attribute of its own and the evolving domain-specific and generally Expert-specific frames of attributes are maintained on the DiscourseStack 123 within the DiscourseHistory object 122.

(EX-1005, [0213]-[0214].) A POSITA would understand that a determination of a confirmation of "'Yes' in response to the system's observation 'So, you're staying in Dublin'" is an interpretation of a second natural language utterance based on a prior first natural language utterance.

220. In order to engage in the dialogue with the user the system must "*determin[e] an interpretation*." O'Neill provides an example dialogue to illustrate this point:

Initial Query

> System: Welcome to the Communicator. How can I help you?
>
> User: I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December.

Notes:

> Referring to FIGS. 2 a and 2 b, key words from the user's utterance like four-star and hotel fall within the accommodation domain, and are provided with a high score by the AccommodationExpert 31. Accordingly the DomainSpotter 60 appoints the AccommodationExpert 31 as the 'handling agent' for the current user enquiry. The AccommodationExpert 31 populates its AccoDialogFrame 131 with whatever relevant information has been supplied by the user.

Implicit Confirmation and Processing of Query

> System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Please wait while I check the database . . .

(EX-1005, [0290]-[0296].)

221. In the above example, the system takes the user's utterance "I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December," and uses the DomainSpotter to determine that the AccommodationExpert is the appropriate domain agent to handle the request, which determines an interpretation so it can provide the response to the user.

222. O'Neill further describes the process the system uses to transfer between agents during an interactive dialogue with a user. (EX-1005, [0261]-[0267].) O'Neill describes recognizing that the user is making a request related to another domain agent from the one it is currently engaged with and deferring that request until later in the conversation. (EX-1005, [0265].) O'Neill then adds the utterance to a UtteranceStore that the system uses to store user utterances. (EX-1005, [0266].) The domain agent then processes the utterances from the UtteranceStore that it can handle or eventually transfers the processing to a different domain agent. (EX-1005, [0266]-[0267].) This further demonstrates the O'Neill system's ability to "*determin[e] the interpretation based*" on the interpretations of domain agents.

223. O'Neill also discloses that "[i]f the DomainSpotter 60 fails to locate a potential handling agent for an 'out-of-domain' utterance in the context of the current service transaction, it polls the other service agents 30 (i.e. does the user want to change from an accommodation enquiry to an enquiry about cinema bookings?)." (EX-1005, [0271].) Therefore, the O'Neill system is continually using the domain agents to process user utterances and "*determining the interpretation based*" on the interpretations of domain agents.

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

| |
|---|
| [1K] determining, by the one or more physical processors, promotional content based on the interpretation; and<br><br>[32K] determine promotional content based on the interpretation; and |

224. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, promotional content based on the interpretation*," and "*determine promotional content based on the interpretation*."

225. Jones discloses: "The system and method disclosed 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].)

226. As discussed above, a POSITA would understand advertisements as a type of "*promotional content*." (*See supra* §V.B.1; *see also* EX-1001, Abstract.) A POSITA would be motivated to use the enhanced interpretation of natural language utterances disclosed in O'Neill to improve the targeting of advertisements disclosed in Jones. Therefore a POSITA would use the targeting of advertisements disclosed in Jones "*based on the interpretation*" from O'Neill.

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

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

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

227. The combination of O'Neill and Jones discloses "*presenting, by the one or more physical processors, the promotional content to a user*."

228. Jones discloses: "The system and method disclosed 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].) Jones further discloses: "Subsequent to selecting 56, process 50 continues presenting, playing or displaying 58 the selected advertisement data to a user. The advertisement data may be displayed while a user is waiting for results of a search." (EX-1006, [0041].) As discussed above in §V.B.1, a POSITA would understand "advertising" to be a type of "*promotional content*."

229. 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].)

## C. Dependent Claim 2

### 1. Utterances Received During Interactive Session [2A]

| |
|---|
| [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: |

230. To the extent that this claim language is limiting, the combination of O'Neill and Jones discloses "*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*," as discussed below in §§V.C.2-3.

231. The system of O'Neill is designed to be "*an interactive session between the user and the computer system*," and to enable a dialogue between the user and the system that evolves dynamically as it progresses:

> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user.

(EX-1005, [0077].)

232. O'Neill also discloses an DiscourseManager that is responsible for handling the "*interactive session between the user and the computer system*":

> The DiscourseManager 22 is responsible for the improved dialogue manager's 20 overall 'conversational style', (i.e. the generic discourse behavior of the improved dialogue manager 20). Given the current technical difficulties with speech recognition and the possibility that a user will misunderstand or change their mind during a dialogue, any system conducting a complex transaction must have a strategy for confirming the semantic contents of a user's utterances.

(EX-1005, [0110].)

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

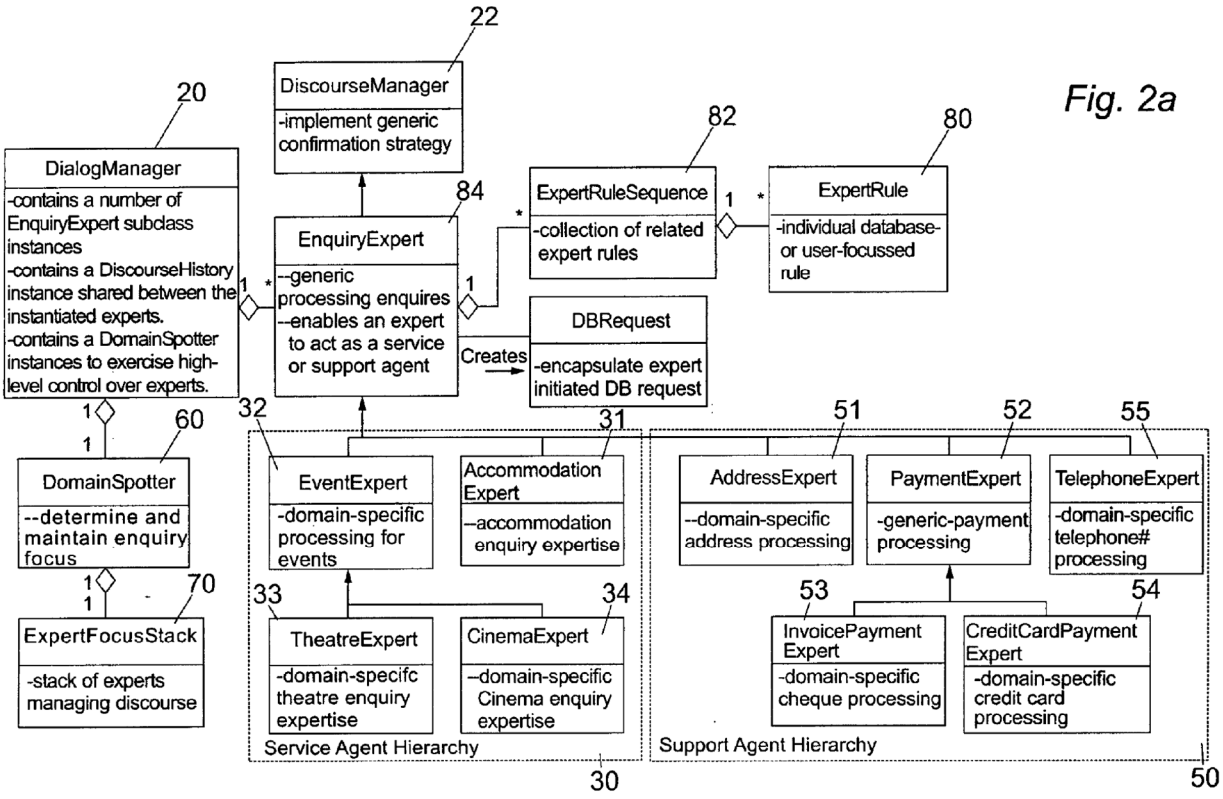
| |
|---|
| [2B] determining, by the one or more physical processors, an activity relating to one or more of the first natural language utterance or the second natural language utterance, |

233. The combination of O'Neill and Jones discloses "*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*."

234. The '536 patent does not define the term "activity." The specification only uses the term briefly in the following context: "Advertisements may be selected based on various criteria, including content of the input (e.g., concepts, semantic indicators, etc.), **an activity related to the input (e.g., a relation to a request, a requested application, etc.)**, user profiles (e.g., demographics, preferences, location, etc.), or in other ways." (EX-1001, 2:31-36.) A POSITA would therefore understand an "activity" to include at utterance that is a "relation to a request" or a

"requested application." A POSITA would understand "relation to a request" includes an utterance that requests the system perform a task would be an "activity."

235. O'Neill discloses determining an "*activity relating to one or more of the first or second natural language utterance*," as the DomainSpotter disclosed in O'Neill determines whether the user is making requests related to certain types of activities, such as hotel accommodations or cinema tickets, before that domain agent then handles the user's request. (EX-1005, Fig. 2a.) This is depicted in Figure 2a:




**O'Neill, Figure 2a**

236. O'Neill provides specific examples of the system "*determining, by the one or more physical processors, an activity relating to one or more*" of the user's

utterances, for example O'Neill describes relating the "*activity*" of booking a cinema ticket to a user utterance:

> On the basis of key words in the user's utterance ('UGC', 'Cinema'), the DomainSpotter 60 recognizes that the user is attempting to change topic from a support agent 50 to a new service agent 30, and at this stage defers the user's request. However, the DomainSpotter 60 places the CinemaExpert 34 at the bottom of the ExpertFocusStack 70 for use later. When the CinemaExpert 34 does eventually become the handling agent, it will be able to make use of the information that the user has already supplied (because this information has been stored in the UtteranceStore 121).
>
> Deferring the User-Initiated Topic Shift
>
> > System: Do you want to enquire about cinema bookings?
> >
> > PROMPT: SPECIFY UserUnderstood QUERY change domain [cinema booking]
> >
> > User: Yes.
> >
> > System: Got it. I'll just complete the payment details first. We'll look at cinema bookings in a moment. Let's go on now with payment details. How would you like to pay?
> >
> > MESSAGE: GIVEN UserUnderstood [true]
> >
> > MESSAGE: transfer not possible from
> >
> > [payment details] defer transfer to
> >
> > [cinema booking]
> >
> > MESSAGE: process payment details
> >
> > PROMPT: SPECIFY PaymentType

(EX-1005, [0394]-[0405]; *see also* EX-1005, [0200]-[0204].)

237. Jones similarly discloses determining "*an activity relating to one or more of the first natural language utterance or the second natural language utterance*," as Jones provides determining a user utterance relates to an activity, such as requesting the system provide information about restaurants:

> For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other 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.

(EX-1006, [0031].)

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

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

238. The combination of O'Neill and Jones discloses "*wherein determining the promotional content comprises determining the promotional content based on the activity*."

239. Jones discloses a system that determines "*the promotional content based on the activity*," because the disclosed system uses a number of factors to weight and determine the appropriate promotional content, including "*activity*":

> An advertisement presented to a user may be based on a weight associated with the advertisement in accordance

> with a single or combined factors such as advertiser contract commitments, bidding price of advertisers, popularity with users, keyword mapping to advertisements, statistical usage (e.g. least recently presented), user demographics, provider choice of advertisement, geographical location of the user and/or provider, etc. For example, a voice query pertaining to a schedule of performances at The Kennedy Center for the Performing Arts in Washington, D.C. from a user who indicates being a resident of New York or whose telephone number corresponds to a New York number, may trigger one or more advertisements related to Broadway shows in New York to be presented to the user, who might respond by purchasing tickets.

(EX-1006, [0029].)

240. In this example, the user is interested in the activity of attending a performance at The Kennedy Center for the Performing Arts and the system determined promotional activity—advertisements related to Broadway shows—based on that activity. (EX-1006, [0029].)

241. Similarly, Jones provides an example in which the user requests information related to restaurants (an activity as discussed above), and then "*determin[es] the promotional content based on the activity*":

> **For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user.** The user telephone number or other 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.

(EX-1006, [0031].)

242. 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: "the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session," which are used by the system to "prepare system utterances in accordance with their generic and domain specific rules." (EX-1005, [0103]-[0104]) Therefore, the O'Neill-Jones combination would use all information when "*determining the promotional content*."

## D. Dependent Claim 3

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

| |
|---|
| [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, |

243. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, context information for the one or more recognized words based on the first natural language utterance*."

244. The '536 patent does not explicitly define the term "context." The '536 patent does provide some disclosure that assist in understanding the scope of the term "context":

> Advertisers may specify various parameters to associate with the advertisements, such as various **contexts or topic concepts (e.g., semantic indicators for a "music" concept may include words such as "music," "tunes," "songs," etc.)**, target demographics (e.g., a preferred audience), marketing criteria or prices for insertion (e.g., dynamic or static pricing based on various marketing criteria), or other information, as would be apparent.

(EX-1001, 2:48-56.)

245. The '536 patent also discloses that context for an utterance is established by having domain agents process the utterance:

> For example, conversational language processor 120 may use context determination module 130 to establish a context for a current utterance by having domain agents 135 competitively generate a context-based interpretation of the utterance (e.g., by scoring possible interpretations and selecting a highest scoring interpretation).

(EX-1001, 4:50-55.)

246. The '536 patent also discloses that context-based interpretation was well known in the art:

> For example, context-based interpretations and responses to a voice-based input may be generated using techniques described in U.S. patent application Ser. No. 11/197,504, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed Aug. 5, 2005, which issued as U.S. Pat. No. 7,640,160 on Dec. 29, 2009, and U.S. patent application Ser. No. 11/212,693, entitled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," filed Aug. 29, 2005, which issued as U.S. Pat. No. 7,949,529 on May 24, 2011, both of which are hereby incorporated by reference in their entirety.

(EX-1001, 4:59-5:3.)

247. O'Neill determines "*context information for the one or more recognized words based on the first natural language utterance*," in the exact same way as disclosed in the '536 patent because O'Neill also determines context by having domain agents "competitively generate a context-based interpretation of the utterance (e.g., by scoring possible interpretations and selecting a highest scoring interpretation)" as discussed above in §V.B.

248. As discussed above, O'Neill discloses that the DomainSpotter determines which domain agent will handle the user request. (§V.B.6.) To make that determination, the DomainSpotter determines the context of the user's enquiry and assigns the handling of the enquiry based on that determination. "For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast." (EX-1005, [0202].)

249. In this example, the DomainSpotter determines "*context information for the one or more recognized words based on the first natural language utterance*" (i.e., that the user is looking for a hotel accommodation in the city of Belfast) and then assigns the AccommodationExpert the responsibility of handling the user's requests because that domain agent has specialized knowledge and processing for that type of request.

250. As discussed above, the DomainSpotter (and the O'Neill system overall) is performed using "*one or more physical processors*." (§V.B.1.)

251. Jones also determines context information for the recognized words. Jones discloses the following example:

> An exemplary process 50 for targeted advertising is illustrated in FIG. 2. As shown in FIG. 2, process 50 begins with associating 52 advertisement data to one or more keywords. As previously discussed, a request may be any type of inquiry or keyword(s) for which a user (i.e., infoseeker™) is seeking specific or general information which may be associated with corresponding advertisement data. For example, advertisement data pertaining to the Indiana Pacers may be associated with the keywords "Reggie Miller."

(EX-1006, [0037.])

252. This example in Jones, in which the system associates the keyword "Reggie Miller" to the context of the "Indiana Pacers" is similar to the context example provided in the '536 patent where the specification associates the keywords "music," "tunes," "songs," to the context of "music." (EX-1001, 2:48-56.)

### 2. Interpretation of the Second Natural Language Utterance Based on Context Information [3B]

| |
|---|
| [3B] wherein the interpretation is determined further based on the context information, |

253. The combination of O'Neill and Jones discloses "*wherein the interpretation is determined further based on the context information*."

254. O'Neill describes that the domain agents use the "*context information*" to "*determine[] the interpretation*" by applying specific "processing rules" to the user's utterances depending on the task the user would like performed:

> According to the invention there is provided a system for conducting a dialogue with a user comprising:
>
> an assignment unit, a plurality of processing
>
> units each of which comprise one or more processing rules, and a plurality of data storage units;
>
> wherein the assignment unit receives a communication from a user, assigns the communication to a processing unit according to the communication's semantic content and stores information relating to the communication in a data storage unit; and
>
> the assigned processing unit processes the communication in accordance with the processing unit's processing rules and provides a response to the user.
>
> …
>
> Desirably, the processing rules of each processing unit reflect the specificities of the tasks they perform.

> Desirably, processing units are hierarchically organized according to the specificity of their processing rules.
>
> Preferably, processing units with more specialized processing rules are substantially independent of those with less specialized processing rules.

(EX-1005, [0023]-[0032].)

255. Claim 26 of O'Neill further describes this process:

> A method of conducting a dialogue with a user comprising the steps of:
>
> > (a) receiving a communication from the user;
> >
> > (b) **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**;
> >
> > (c) storing information from the communication in one or more data-storage units;
> >
> > (d) **using one or more of the processing rules of the assigned processing unit to process the communication**;
> >
> > (e) forwarding a response to the user; and
> >
> > (f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.

(EX-1005, claim 26.)

256. O'Neill provides an example in which a user is looking for a hotel accommodation and the system assigns the request for the AccomodationExpert to

process the request, which uses domain specific information and context from the user's request to process the request:

> Referring to FIGS. 2 a and 2 b, key words from the user's utterance like four-star and hotel fall within the accommodation domain, and are provided with a high score by the AccommodationExpert 31. Accordingly the DomainSpotter 60 appoints the AccommodationExpert 31 as the 'handling agent' for the current user enquiry. The AccommodationExpert 31 populates its AccoDialogFrame 131 with whatever relevant information has been supplied by the user.
>
> Implicit Confirmation and Processing of Query
>
> System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Please wait while I check the database…

(EX-1005, [0294]-[0296].)

257. O'Neill also discloses "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].) 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].)

258. As I discuss above, §IV.A, the '536 patent states that "context-based interpretations and responses to a voice-based input" was known in the art and not a novel feature of the alleged invention. (EX-1001, 4:59-5:3.)

### E. Dependent Claim 4

> [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.

259. The combination of O'Neill and Jones discloses "*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*."

260. O'Neill discloses that once each of the domain agents, which it calls "service agents," receives and processes the second natural language utterance (which it receives from the natural language parser), each domain agent interprets the utterance and provides a score of the utterance's relevance to its domain:

> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60.
>
> The service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific

> heuristics to the more detailed processing of the user's enquiry. For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast.

(EX-1005, [0201]-[0202].)

261. As discussed above in §V.B.10, O'Neill further describes the process the system uses to transfer between agents during an interactive dialogue with a user, which demonstrates that the O'Neill system has each domain agent determine an interpretation of the utterance to determine which agent is most appropriate to handle the user utterance. (EX-1005, [0261]-[0267]; *see also* EX-1005, [0271] (describing polling other domain agents to interpret request).)

## F. Dependent Claim 5

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

| |
|---|
| [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<br><br>[5B] associating, by the one or more physical processors, the second interpretation with a second interpretation score, |

262. The combination of O'Neill and Jones discloses "*further comprising, associating, by the one or more physical processors, the first interpretation with a first interpretation score*," and "*associating, by the one or more physical processors, the second interpretation with a second interpretation score*."

263. As discussed above, O'Neill discloses that once each of the domain agents, which it calls "service agents," receives and processes the second natural language utterance (which it receives from the natural language parser), each domain agent interprets the utterance and provides an "*interpretation score*" of the utterance's relevance to its domain:

> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60.
>
> The service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific heuristics to the more detailed processing of the user's enquiry. For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast.

(EX-1005, [0201]-[0202].)

264. O'Neill provides the following example of two domain agents interpreting an utterance and associating it with an "*interpretation score*" of the utterance's relevance to its domain:

> In addition, specialized agents also attain a higher score for specialized parser tags than generic agents. For example, a user request such as "I'd like to go to see THE LION KING 2." might parse as event_enquiry: [Eventtype][Movies].THE LION KING 2.
>
> In this case, although the EventExpert 32 could award a score for the event_enquiry, the CinemaExpert 34, as a child of EventExpert 32, would award a score not only for the event enquiry, but for Movies as well, and so would be the winner.

(EX-1005, [0203]-[0204].)

265. In this example, the EventExpert and CinemaExpert both interpreted the user's utterance and associated a score with its interpretation. Specifically, the CinemaExpert scored higher because "I'd like to go see THE LION KING 2" is interpreted as a request related to movies and assigned to the CinemaExpert domain agent. (EX-1005, [0203]-[0204].)

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

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

266. The combination of O'Neill and Jones discloses "*wherein determining the interpretation comprises determining the first interpretation as the interpretation based on the first interpretation score and the second interpretation score*."

267. As discussed above, O'Neill discloses that "[t]he service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the

user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific heuristics to the more detailed processing of the user's enquiry." (EX-1005, [0202].) Once the domain agent (service agent) is selected, that domain agent is responsible for handling the user's request.

268. O'Neill explains that, in the example discussed above, when the user makes the request "'I'd like t ogo see THE LION KING 2,'" the system "might parse as event_enquiry: [Eventtype] [Movies].THE LION KING 2." (EX-1005, [203].) In this situation, "although the EventExpert 32 could award a score for event_enquiry, the CinemaExpert 34, as a child of EventExpert 32, would award a score not only for the event enquiry, but for Movies as well, so would be the winner." (EX-1005, [0204].)

269. Therefore, O'Neill discloses "*determining the interpretation comprises determining the first interpretation as the interpretation based on the first interpretation score and the second interpretation score*" because comparing the scores of the service agents is determining the interpretation of the utterance based on all of the scores the DomainSpotter is provided.

### G. Dependent Claim 6

| |
|---|
| [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. |

270. The combination of O'Neill and Jones discloses "*wherein the first interpretation is determined as the interpretation based on the first interpretation score being greater than the second interpretation score*."

271. As discussed above, O'Neill discloses the scores returned by each domain agent are compared and the domain agent with the highest score is selected to handle the user's request:

> Each service agent 30 is provided with a range of integer values for scoring the degree of relevance it assigns to different domain-specific parse tags. Accordingly, each service agent 30 scores the parse of the initial user utterance against the semantic categories that it can process and returns the score to the DomainSpotter 60.
>
> The service agent 30 that attains the highest score is deemed to be the most appropriate handling agent for the user's enquiry by the DomainSpotter 60. Accordingly, this service agent 30 is selected to apply its domain-specific heuristics to the more detailed processing of the user's enquiry. For example, the AccommodationExpert 31 might score highest and so become the handling agent if the user had been asking about hotels in Belfast.

(EX-1005, [0201]-[0202].)

### H. Dependent Claims 7 and 51

272. Dependent claims 7 and 51 each add the exact same limitation over the independent claims each depends from and I therefore have analyzed them together.

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

[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

[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

273. The combination of O'Neill and Jones discloses "*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*."

274. As discussed above, O'Neill discloses a system that engages in a "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":

> a method of conducting a dialogue with a user comprising the steps of:
>
> > (a) receiving a communication from the user;
> >
> > (b) 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;
> >
> > (c) storing information from the communication in one or more data-storage units;
> >
> > (d) using one or more of the processing rules of the assigned processing unit to process the communication;
> >
> > (e) forwarding a response to the user; and
> >
> > **(f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.**

(EX-1005, [0052]-[0058]; *see also* EX-1005, [0023]-[0027].)

275. O'Neill also provides an example of a user making three utterances to book a hotel accommodation:

> In this example, a user wishes to book a four star hotel in Belfast from 15-20 December. When the automatic dialogue system returns a list of four start hotels, the user changes the request to a three star hotel. The automatic dialogue system returns a list of three star hotels and the user selects one and asks to book a double room. The user then asks for a single room rather than a double.

(EX-1005, [0287].)

276. O'Neill provides specific examples of dialogue between the user and the system that demonstrates that the system discloses receiving "*a third natural language utterance relating to one or more of the first natural language utterance or the second natural language utterance*":

> System: Welcome to the Communicator. How can I help you?
>
> User: I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December.
>
> …
>
> System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Please wait while I check the database ...
>
> …
>
> User: Could you make that a three-star?
>
> …
>
> System: OK, got you. That's three-star accommodation. Is that OK?
>
> …
>
> User: Yes, that's right.
>
> System: Please wait while I check the database.

(EX-1005, [0291]-[0321].)

277. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (*Supra* §V.B.1.)

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

| |
|---|
| [7B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content<br><br>[51B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content. |

278. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content*."

279. O'Neill discloses "*determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content*," because the system disclosed by O'Neill uses all information from the interactive dialogue session with the user, including the promotional content when combined with Jones, to determine the interpretation and response to the user's enquiries, including the "*third natural language utterance*." (*See* §V.B.)

280. This is evidenced by the example dialogue above in §V.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 of "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*."

281. O'Neill discloses that the dialogue manager continues to use information from prior parts of the conversation with the user to interpret and respond to the user's current enquiry:

> Preferably, **the method includes a step of retrieving information from a previous communication stored in one or more of the data-storage units** and combining the information with information from the current communication to enable the completion of a task associated with the previous communication.

(EX-1005, [0062].)

282. O'Neill further discloses that it retrieves information collected earlier in the conversation to assist interpreting the current utterance from the user:

> Furthermore, the improved dialogue manager provides a facility whereby the dialogue manager may retrieve information collected at an earlier point in a dialogue to assist in a current query.
>
> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse

> forward, and across domain boundaries if required by the user.

(EX-1005, [0076]-[0077].)

283. O'Neill describes the collection and use of information within a Knowledge Hierarchy, including information that is generated by the system, like promotional content, and provided to the user:

> **Referring to FIG. 2 b, the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session**. In particular, the knowledge-storing components store data values elicited by the system from the user, or inferred by the system itself, in the course of a dialogue. These components can be collectively described as the Knowledge Hierarchy.
>
> The components in the Knowledge Hierarchy also maintain information that indicates how confirmed each datum is, and what the system intends to do (if anything) to confirm that datum adequately. The data values stored in the knowledge-storing elements are the system's discourse knowledge, and are used by components of the Expertise Hierarchy as they prepare system utterances in accordance with their generic and domain specific rules.

(EX-1005, [0103]-[0104].)

284. O'Neill further discloses:

> The above-described information-storage mechanism enables the improved dialogue manager **20** to store the dialogue product (i.e. the information elicited during the evolving dialogue of a dialogue session) in a tree-like structure (henceforth known as the knowledge tree). The tree-like structure of the dialogue product emerges from the fact that a dialogue frame may include as part of its Attributes, links to dialogue frames of other types (e.g.

> TheatreDialogFrame 133 may include a link to a PaymentDialogFrame 152).

(EX-1005, [0184].)

285. Therefore, O'Neill discloses interpretation of the "third natural language utterance"—and all natural language utterances from the user—based on promotional content because O'Neill's dialogue manager retrieves and uses all types of information used in earlier points in the conversation with the user.

286. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (§V.B.1.)

## I. Dependent Claim 8

| |
|---|
| [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. |

287. The combination of O'Neill and Jones discloses "*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*."

288. As discussed above, O'Neill discloses a system that engages in a "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":

> [A] method of conducting a dialogue with a user comprising the steps of:
> (a) receiving a communication from the user;
> (b) 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;
> (c) storing information from the communication in one or more data-storage units;
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
> (e) forwarding a response to the user; and
> **(f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.**

(EX-1005, [0052]-[0058]; *see also* EX-1005, [0023]-[0027].)

289. O'Neill also provides an example of a user making three utterances to book a hotel accommodation "*received during an interactive session between the user and the computer system*":

> In this example, a user wishes to book a four star hotel in Belfast from 15-20 December. When the automatic dialogue system returns a list of four start hotels, the user changes the request to a three star hotel. The automatic dialogue system returns a list of three star hotels and the user selects one and asks to book a double room. The user then asks for a single room rather than a double.

(EX-1005, [0287].)

### J. Dependent Claim 9

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

290. The combination of O'Neill and Jones discloses "*wherein the third natural language utterance is received after the second natural language utterance*."

291. As discussed above, O'Neill discloses a system that engages in a "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" "*wherein the third natural language utterance is received after the second natural language utterance*":

> [A] method of conducting a dialogue with a user comprising the steps of:
> (a) receiving a communication from the user;
> (b) 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;
> (c) storing information from the communication in one or more data-storage units;
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
> (e) forwarding a response to the user; and
> **(f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated.**

(EX-1005, [0052]-[0058]; *see also* EX-1005, [0023]-[0027].)

292. O'Neill also provides an example of a user making three utterances to book a hotel accommodation with a third utterance coming after the second utterance:

> In this example, a user wishes to book a four star hotel in Belfast from 15-20 December. When the automatic dialogue system returns a list of four start hotels, the user changes the request to a three star hotel. The automatic dialogue system returns a list of three star hotels and the user selects one and asks to book a double room. The user then asks for a single room rather than a double.

(EX-1005, [0287]; *see also* EX-1005, [0291]-[0321] (providing another example of a user booking a hotel accommodation with three utterances).)

## K. Dependent Claim 10

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

| |
|---|
| [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; |

293. The combination of O'Neill and Jones discloses "*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*."

294. As discussed above, O'Neill discloses a system that engages in a "dialogue" between a user and the system that repeats under the user stops the dialogue, which includes "third natural language utterances":

> "[A] method of conducting a dialogue with a user comprising the steps of:
>
> (a) receiving a communication from the user;
>
> (b) 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;
>
> (c) storing information from the communication in one or more data-storage units;
>
> (d) using one or more of the processing rules of the assigned processing unit to process the communication;
>
> (e) forwarding a response to the user; and
>
> (f) repeating the above steps until the user fails to issue further communications or indicates that the dialogue is terminated."

(EX-1005, [0052]-[0058]; *see also* EX-1005, [0023]-[0027].)

295. O'Neill further discloses identifying "*reference words*" which O'Neill calls "Keywords" that determine which path should be used:

> For example, at the outset of an accommodation enquiry, the service dialogue frame for the enquiry (i.e. AccoDialogFrame 131) will not generally contain an explicit link to a payment dialogue frame (i.e. PaymentDialogFrame 152). However, since the DomainSpotter 60 can determine which agents provide payment support, the improved dialogue manager 20 can generate a number of potential discourse paths relating to payment. **Keywords in the user's utterance determine which path is in fact used and which payment-related dialogue frames are explicitly linked to the accommodation dialogue frame**.

(EX-1005, [0254], [0140] (recognizing reference words such as "four start" and "hotel" associated with booking a hotel reservation).)

296. Also, as discussed above, the natural language parser used in the O'Neill system outputs keywords and phrases: "The improved dialogue manager accepts as input semantically tagged key-words and phrases from the Phoenix natural language parser (W. Ward, Proceedings of ICSLP 94, pp. 83-86, Yokohama, 1994). The improved dialogue manager is interfaced to the Galaxy hub 14 through a Java DialogServer class which includes the relevant hub package (importgalaxy.server.*;) and creates and populates frames of information in the Galaxy format." (EX-1005, [0278].) As discussed above in §V.B.5, O'Neill discloses an "automatic speech recognizer" (EX-1005, [0013], Fig. 1, [0096].)

297. 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, [0140]; *see also* EX-1005, [0291]-[0320].) In the example discussed above in §V.H.1, the user confirms to the system that the user wants a three-star hotel in the "*third natural language utterance*." (EX-1005, [0287]; *see also* EX-1005, [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 discuss below. For example, the

word "that" may be a reference word recognized by the system as referencing the three-star hotel in the prior dialogue. .

298. Further, as discussed above, Jones also discloses a system that provides "targeted advertising" "using voice recognition by identifying a keyword in a spoken phrase, looking up at the keyword in an index having corresponding advertising and providing the advertising." (EX-1006, [0066].)

299. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (§V.B.1.)

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

| |
|---|
| [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; |

300. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, that the one or more reference words refer to an item associated with the promotional content*."

301. The '536 patent does not define the term "item." In fact, the specification of the '536 patent does not use the term "item" a single time. The claims of the '536 patent. 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."

Therefore, I understand item can refer to tangible goods and electronic goods.

302. Jones discloses:

> The system and method disclosed 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].)

303. Jones also discloses:

> Targeted advertisement is provided to mobile device users based on one or more keywords in words spoken by the mobile device users. The users may submit voice requests using mobile phones or other mobile networked devices, relevant keyword(s) are determined from the requests and corresponding advertisements and/or products associated with the keyword(s) are provided to the users while the users are waiting for responses to the requests.

(EX-1006, Abstract.)

304. Jones also explains:

> After transcribing 54 keyword(s), process 50 continues to selecting 56 advertisement data for the transcribed keyword(s). For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes.

(EX-1006, [0040].)

305. As discussed above, the O'Neill system also determines that one or more reference words refer to an item. In the example discussed above, after the user refers to the Lion King 2, the system determines that the user is referring to purchasing a movie ticket, an "*item*." Therefore, a POSA would combine these teachings in O'Neill with the targeted advertisement teachings in Jones to further enhance the selection of promotional content.

306. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*determining, by the one or more physical processors, that the one or more reference words refer to an item associated with the promotional content*." (§V.B.1.)

### 3. Initiating a Purchase Transaction [10C]

| |
|---|
| [10C] 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. |

307. The combination of O'Neill and Jones discloses "*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*."

308. O'Neill further discloses initiating a purchase transaction after identifying "reference words" which O'Neill calls "Keywords":

> For example, at the outset of an accommodation enquiry, the service dialogue frame for the enquiry (i.e. AccoDialogFrame 131) will not generally contain an explicit link to a payment dialogue frame (i.e. PaymentDialogFrame 152). However, since the DomainSpotter 60 can determine which agents provide payment support, the improved dialogue manager 20 can generate a number of potential discourse paths relating to payment. **Keywords in the user's utterance determine which path is in fact used and which payment-related dialogue frames are explicitly linked to the accommodation dialogue frame**.

(EX-1005, [0254].)

309. O'Neill also discloses initiating a purchase transaction associated with the purchase of a cinema ticket. (*See* EX-1005, [0366]-[0389].)

310. Jones similarly discloses the ability to initiate "*a purchase transaction associated with the item based on the determination that the one or more reference words refer to the item*":

> Subsequent to delivering 74 the advertisement, the process 70 moves to providing 76 an option to receive a coupon for the goods/services indicated in the advertisement. Using the same example in the previous discussion, if a user interacts with the advertisement pertaining to the specific discounted Dell products, an option to receive a coupon for the products may be provided to the user (see also FIG. 3B). In the case of electronic merchandise (eg. MP3 downloads) a product may be delivered in addition to an advertisement. For example the advertisement may contain video or audio 'clips' ad the consumer may elect to purchase the full-length product, which is then supplied to him.

(EX-1006, [0055].)

311. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*determining, by the one or more physical processors, that the one or more reference words refer to an item associated with the promotional content*." (§V.B.1.)

### L. Dependent Claim 11

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

312. The combination of O'Neill and Jones discloses "*wherein the item includes a purchasable good*."

313. As discussed above in §V.K.2, Jones discloses the ability to initiate "*a purchase transaction*" of an item that is a "purchasable good*"* like an MP3 song:

> Subsequent to delivering 74 the advertisement, the process 70 moves to providing 76 an option to receive a coupon for the goods/services indicated in the advertisement. Using the same example in the previous discussion, if a user interacts with the advertisement pertaining to the specific discounted Dell products, an option to receive a coupon for the products may be provided to the user (see also FIG. 3B). In the case of electronic merchandise (eg. MP3 downloads) a product may be delivered in addition to an advertisement. For example the advertisement may contain video or audio 'clips' ad the consumer may elect to purchase the full-length product, which is then supplied to him.

(EX-1006, [0055].)

314. Jones also provides examples of other types of advertisements that can be provided to the user that include purchasable goods, like iPods:

> After transcribing 54 keyword(s), process 50 continues to selecting 56 advertisement data for the transcribed keyword(s). For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes.

(EX-1006, [0040].)

### M. Dependents Claims 12 and 52

> [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.
>
> [52] The method of claim 51, 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.

315. Claim 12 and 52 depend from claims 7 and 51, respectively, and each add the limitation "*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*."

316. The combination of O'Neill and Jones discloses "*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*."

317. Specifically, Jones discloses that "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, the database 18 may be updated to reflect additional ad revenue with a credit of points and/or compensation, if appropriate, to the provider (or searcher).” (EX-1006, [0028].)

318. Jones also provides examples in which the system provides the customer coupons to use in the purchase of an item or allows the user to purchase the item immediately:

> Subsequent to delivering 74 the advertisement, the process 70 moves to providing 76 an option to receive a coupon for the goods/services indicated in the advertisement. Using the same example in the previous discussion, if a user interacts with the advertisement pertaining to the specific discounted Dell products, an option to receive a coupon for the products may be provided to the user (see also FIG. 3B). In the case of electronic merchandise (eg. MP3 downloads) a product may be delivered in addition to an advertisement. For example the advertisement may contain video or audio ‘clips’ ad the consumer may elect to purchase the full-length product, which is then supplied to him.

(EX-1006, [0055].)

319. O’Neill provides the framework to initiate and complete the purchase transaction described in Jones inside the interactive dialogue system. O’Neill discloses that the system can initiate and complete a purchase transaction inside the dialogue with the user:

> For example, at the outset of an accommodation enquiry, the service dialogue frame for the enquiry (i.e.

> AccoDialogFrame 131) will not generally contain an explicit link to a payment dialogue frame (i.e. PaymentDialogFrame 152). However, since the DomainSpotter 60 can determine which agents provide payment support, the improved dialogue manager 20 can generate a number of potential discourse paths relating to payment. **Keywords in the user's utterance determine which path is in fact used and which payment-related dialogue frames are explicitly linked to the accommodation dialogue frame.**

(EX-1005, [0243].)

320. O'Neill further discloses: "Preferably, the processing units includes at least one processing unit adapted to acquire payment details from a user." (EX-1005, [0068].)

321. O'Neill discloses that certain domain agents can be responsible for handling acquiring the user's address and payment details for the purchase:

> In the present example, the class of service agents 30 include an AccommodationExpert 31 and an EventExpert 32 that are respectively responsible for searching for and booking accommodation and event tickets. In addition, the class of support agents 50 include an AddressExpert 51 and a PaymentExpert 52 that are respectively responsible for acquiring the user's address and payment details. It 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, [0123].)

322. A POSITA would understand that the targeted advertising disclosed in Jones can be combined with the interactive dialogue system disclosed in O'Neill and

that the combined system could initiate "*a purchase transaction associated with the promotional content*."

323. Also, as discussed above in V.B.11, promotional content can be selected and provided to the user based on all of the user's utterance as the system in O'Neill processes all of the user's utterances to formulate its response to the user, including providing promotional content.

324. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (§V.B.1.)

## N. Dependent Claim 16

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

| |
|---|
| [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, |

325. The combination of O'Neill and Jones discloses "*identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance*."

326. O'Neill discloses identifying one or more requests related to one or more of the first or second natural language utterance, as the system disclosed in

O'Neill identifies whether the user is making requests for certain information or requests to assist with certain tasks, such as requesting assistance booking hotels, or event tickets such as theatre tickets or cinema tickets. (EX-1005, Fig. 2a.)

327. O'Neill provides specific examples of the system determining that the user's utterance relate to a specific request, like booking a cinema ticket:

> On the basis of key words in the user's utterance ('UGC', 'Cinema'), the DomainSpotter 60 recognizes that the user is attempting to change topic from a support agent 50 to a new service agent 30, and at this stage defers the user's request. However, the DomainSpotter 60 places the CinemaExpert 34 at the bottom of the ExpertFocusStack 70 for use later. When the CinemaExpert 34 does eventually become the handling agent, it will be able to make use of the information that the user has already supplied (because this information has been stored in the UtteranceStore 121).
>
> Deferring the User-Initiated Topic Shift
>
> System: Do you want to enquire about cinema bookings?
>
> PROMPT: SPECIFY UserUnderstood QUERY change domain [cinema booking]
>
> User: Yes.
>
> System: Got it. I'll just complete the payment details first. We'll look at cinema bookings in a moment. Let's go on now with payment details. How would you like to pay?
>
> MESSAGE: GIVEN UserUnderstood [true]
>
> MESSAGE: transfer not possible from
>
> [payment details] defer transfer to
>
> [cinema booking]
>
> MESSAGE: process payment details
>
> PROMPT: SPECIFY PaymentType

(EX-1005, [0394]-[0405].)

328. O'Neill a further example where a request is associated with the user's second natural language utterance (i.e., request for a three-star hotel):

> [0291] System: Welcome to the Communicator. How can I help you?
> [0292] User: I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December.
> …
> [0296] System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Please wait while I check the database ...
> …
> [0312] **User: Could you make that a three-star?**
> …
> [0314] System: OK, got you. That's three-star accommodation. Is that OK?
> …
> [0320] User: Yes, that's right.
> [0321] System: Please wait while I check the database.

(EX-1005, [0291]-[0321].)

329. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance*." (§V.B.1.)

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

| |
|---|
| [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. |

330. The combination of O'Neill and Jones discloses "*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*."

331. Specifically, Jones provides examples of providing promotional content that relates to one or more requests made by the user:

> After transcribing 54 keyword(s), process 50 continues to selecting 56 advertisement data for the transcribed keyword(s). For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes.

(EX-1006, [0040].)

332. Jones provides additional examples of obtaining promotional content that relates to the user's request:

> For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other 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.

(EX-1006, [0031].)

333. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (§V.B.1.)

## O. Dependent Claim 17

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

| |
|---|
| [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 |

334. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, one or more activities in the one or more requests*."

335. As discussed above, (§V.N.2), O'Neill discloses a system that determines one or more activities from the user's requests. For example, if the user makes a request that the system book a hotel accommodation for the user, the system disclosed in O'Neill will then determine that the user is instructing the system to assist the user in the activity of booking a hotel accommodation. (EX-1005, [0274]-[0342].) As another example, O'Neill discloses that the system can determine based on the user's request that the user would like assistance booking a cinema ticket. The system will then assist the user with activity of booking the cinema ticket. (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].)

336. As discussed above, O'Neill and Jones both perform the disclosed methods using "*one or more physical processors*," including "*determining, by the one or more physical processors, one or more activities in the one or more requests*." (§V.B.1.)

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

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

337. The combination of O'Neill and Jones discloses "*performing, by the one or more physical processors, the one or more activities*."

338. As discussed above, §V.O.1, the system disclosed by O'Neill performs the activity for the user after it determines the activity based on the user's request. For example, the system will perform the activity of booking the user a hotel accommodation or perform the activity of booking the user a cinema ticket by collecting payment information from the user and other necessary information, like the user's address. (EX-1005, [0391]-[0403].) O'Neill further assists the user with payment by collecting payment information from the user:

> System: Yes, that's reserved. Let's go on now with payment details. Let's start with the address. Can you tell me your house number please?
> . . .
> User: It's 3 Malone Road, Belfast BT7
> . . .
> System: OK, I've got all your address details. Let's go on now with payment details. How would you like to pay?

(EX-1005, [0366]-[0370].)

339. Jones also discloses determining that a user's request relates to the activity of obtaining information regarding certain restaurants the user is interested in:

> For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other 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.

(EX-1006, [0031].)

340. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*performing, by the one or more physical processors, the one or more activities*." (§V.B.1.)

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

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

341. The combination of O'Neill and Jones discloses "*wherein determining the promotional content comprises determining the promotional content based on the one or more activities*."

342. The elements in limitation [17C] are substantively identical to those in limitation [2C] above in §V.C.3, the O'Neill-Jones combination discloses this limitation. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [2C] wherein determining the promotional content comprises determining the promotional content based on the activity. | [17C] wherein determining the promotional content comprises determining the promotional content based on the one or more activities. |

343. For the same reasons discussed above in §V.C.3 for claim limitation [2C], the combination of O'Neill and Jones discloses this limitation.

### P. Dependent Claim 18

> [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.

344. The combination of O'Neill and Jones discloses "*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*."

345. Jones discloses "*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*," as Jones discloses a system where a user can contact the system to obtain information about a business, including the phone number:

> For example, the consumer 62 calls an 800 number with a query about a business in a particular locality and process 60 captures subscriber identity information such as a phone number of the consumer.

(EX-1006, [0043].)

346. The system discloses that it also has the ability to connect a user directly to a business:

> Further, a user may be transferred directly to a particular business if the user presses a key or button during an advertisement.

(EX-1006, [0062].)

347. The system also determines the promotional content based on the user's request for information about the business, including the phone number:

> For example, a user searching for a location of a particular restaurant in Carmel, Ind., may be presented with

> advertisements pertaining to popular attractions in Carmel while the location of the particular restaurant is being searched.

(EX-1006, [0030].)

### Q. Dependent Claim 19

> [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.

348. The combination of O'Neill and Jones discloses "*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*."

349. O'Neill discloses that the system disclosed can be used and incorporated with route guidance systems, which provide a "navigation route" as claimed:

> Reference will now be made in detail to the preferred embodiment of the present invention, example of which is illustrated in the accompanying drawings. The improved dialogue manager will be described by way of example in an accommodation and event booking and payment system. **It will be appreciated 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 or other transactional systems.**

(EX-1005, [0089].) A POSITA would know that prior art "route-guidance systems" disclosed in O'Neill are designed to "*calculat[e] the navigation route for the user*." As discussed above, the interactive dialogue system of O'Neill uses all of the user's utterances to determine the system's responses to the user. In the combination of O'Neill and Jones, the system would include all user utterances, including utterances related to navigation routes to determine the promotional content to provide the user.

350. While Jones does not disclose calculating a navigation route, it does disclose examples where users request information regarding a particular location:

> Users of the system 10 may be provided with an option of submitting a request directly to resource(s) identified in the database 18 based on keywords or categories extracted from the request, when a user opts to conduct a search without necessarily invoking assistance from a provider. For example, a user searching for a location of a particular restaurant in Carmel, Ind., may be presented with advertisements pertaining to popular attractions in Carmel while the location of the particular restaurant is being searched.

(EX-1006, [0030].)

351. Moreover, Jones discloses using the user's location to determine what promotional content to provide the user:

> The user telephone number or other 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. Geographic information may also be requested by the system during the call, usually during the beginning of the session. The request from the system (or guide) may be of the form of voice or text. The response by the user may also be in the form of voice, text, or touchtone entry. The guide may ask the user or an automated voice may ask the user for location information. The location information may be a zip code, a street number, a street intersection, a point of interest, latitude and longitude coordinates, an area code, or any other form of location information. This location information, once gathered by the system, may then be stored. Keywords may be linked to guides and/or advertisements that are tied to a specific geographic area that has a well-defined boundary. Alternatively, a guide or advertisement may be chosen based on the "closest" guide or advertiser to the user who is making the query.

(EX-1006, [0031].)

352. Therefore, a POSITA would understand that Jones teaches using a user's location to determine promotional content and could similarly use a user's navigation route to determine promotional content.

## R. Dependent Claim 20

> [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.

353. The combination of O'Neill and Jones discloses "*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*."

354. O'Neill discloses that the system can be used broadly for other automatic dialogue systems, one of which could be processing user requests to be presented with media:

> Reference will now be made in detail to the preferred embodiment of the present invention, example of which is illustrated in the accompanying drawings. The improved dialogue manager will be described by way of example in an accommodation and event booking and payment system. **It will be appreciated 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 or other transactional systems.**

(EX-1005, [0089].)

355. 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. The system also discloses providing a "advertisement" along with the delivery of the "*media item*":

> Subsequent to delivering 74 the advertisement, the process 70 moves to providing 76 an option to receive a coupon for the goods/services indicated in the advertisement. Using the same example in the previous discussion, if a user interacts with the advertisement pertaining to the specific discounted Dell products, an option to receive a coupon for the products may be provided to the user (see also FIG. 3B). **In the case of electronic merchandise (eg. MP3 downloads) a product may be delivered in addition to an advertisement. For example the advertisement may contain video or audio 'clips' ad the consumer may elect to purchase the full-length product, which is then supplied to him.**

(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*."

356. As discussed above, 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.

## S. Dependent Claim 21

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

| |
|---|
| [21A] 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 |

357. The combination of O'Neill and Jones discloses "*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*."

358. As discussed above, (§V.N.2), O'Neill discloses a system that determines one or more activities from the user's requests. For example, if the user makes a request that the system book a hotel accommodation for the user, the system disclosed in O'Neill will then determine that the user is instructing the system to assist the user in the activity of booking a hotel accommodation. (EX-1005, [0274]-[0342].)

359. As another example, O'Neill discloses that the system can determine based on the user's request that the user would like assistance booking a cinema ticket. The system will then assist the user with activity of booking the cinema ticket. (EX-1005, [0391]-[0403].)

360. Jones also discloses assisting users with location restaurants, which may include assisting the user to make a reservation:

> For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other 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.

(EX-1006, [0031].)

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

| |
|---|
| [21B] wherein determining the promotional content comprises determining the promotional content based on information relating to the reservation. |

361. The combination of O'Neill and Jones discloses "*determining the promotional content comprises determining the promotional content based on information relating to the reservation*."

362. Similarly, Jones discloses determining and presenting users with targeted promotional content based on one or more activities. For example, Jones discloses a system that provides users information related to restaurants in a location:

> For example, a user searching for a location of a particular restaurant in Carmel, Ind., may be presented with advertisements pertaining to popular attractions in Carmel while the location of the particular restaurant is being searched.

(EX-1006, [0030].)

363. A POSITA would therefore combine the targeted advertising related to restaurants with the booking features disclosed by O'Neill. The O'Neill system is collecting information about the users intent to make certain purchases related to those activities, i.e., paying for a hotel room or paying for a cinema ticket. (EX-1005, [0391]-[0403].) Those activities provide 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 related to those activities.

## T. Dependent Claim 22

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

| |
|---|
| [22A] 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 |

364. The combination of O'Neill and Jones discloses "*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*."

365. As discussed above, (§V.S.1), O'Neill discloses a system that determines one or more activities from the user's requests. For example, if the user makes a request that the system book a hotel accommodation for the user, the system disclosed in O'Neill will then determine that the user is instructing the system to

assist the user in the activity of booking a hotel accommodation. (EX-1005, [0274]-[0342].)

366. As another example, O'Neill discloses that the system can determine based on the user's request that the user would like assistance booking a cinema ticket. The system will then assist the user with activity of booking the cinema ticket. (EX-1005, [0391]-[408].)

367. O'Neill discloses that the system can provide alternative recommendations for bookings (i.e., reservations). For example, the system in O'Neill 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:

> System: Welcome to the Communicator. How can I help you?
>
> User: I'd like to book a four-star hotel in Belfast from the fifteenth of December to the twentieth of December.
>
> …
>
> System: So that's an accommodation booking for a four-star hotel in Belfast from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. 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: Could you make that a three-star?
>
> …

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

…

User: Yes, that's right.

System: Please wait while I check the database

…

System: You can choose between the following hotels: Jurys and The Lisdara

…

User: I'll go for a double room at Jury's

…

System: So, let me just check that. You want a double room at the three-star Jury's hotel in Belfast, from the fifteenth of December two thousand and three to the twentieth of December two thousand and three. Are those details correct?

(EX-1005, [0291]-[0327].)

368. Jones also discloses:

> For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user. The user telephone number or other 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.

(EX-1006, [0031].)

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

| |
|---|
| [22B] 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. |

369. The combination of O'Neill and Jones discloses "*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*."

370. O'Neill also discloses assisting the user in the same conversation with changing the room reservation from a double room to a single room, which is an alternative to the first reservation the system presented. (EX-1005, [0290]-[0360].)

371. O'Neill discloses a system that assists users with various activities, such as booking a hotel accommodation or booking a cinema ticket, as discussed above. (EX-1005, [0274]-[0342], [0391]-[0408].) The O'Neill system is therefore collecting information about the users intent to make certain purchases related to those activities, i.e., paying for a hotel room or paying for a cinema ticket. Those activities provide an opportunity for the system to present the user with promotional content or advertisements related to those activities, including the second reservation that is an alternative to the first reservation.

372. Similarly, Jones discloses user may be provided promotional content relating to businesses that are alternatives from the businesses they are searching for:

> For example, a user searching for a location of a particular restaurant in Carmel, Ind., may be presented with advertisements pertaining to popular attractions in Carmel while the location of the particular restaurant is being searched.

(EX-1006, [0030].)

373. A POSITA would understand from this example that the system in Jones could be used to present users with promotional content pertaining to an alternative restaurant reservation from the first reservation the user is searching for.

374. As discussed above §V.B.11, the combination of O'Neill and Jones would use all relevant information—including "*information relating to a second reservation that is an alternative to the first reservation*"—to determine "*the promotional content*."

## U. Dependent Claim 23

### 1. Receiving a Third Natural Language Utterance [23A]

| |
|---|
| [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; |

375. The combination of O'Neill and Jones discloses "*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*."

376. This limitation of claim 23 is substantially the same limitation that is addressed in §V.H.1 regarding claim limitation [7A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

377. For the same reasons discussed above in §V.H.1 for claim limitation [7A], the combination of O'Neill and Jones discloses this limitation.

### 2. Determining an Interpretation of Third Natural Language Utterance Based on Promotional Content [23B]

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

378. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content*."

379. This limitation of claim 23 is substantially the same limitation that is addressed in §V.H.2 regarding claim limitation [7B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [7B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content | [23B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content; and |

380. For the same reasons discussed above in §V.H.2 for claim limitation [7B], the combination of O'Neill and Jones discloses this limitation.

### 3. Initiating a Purchase Transaction [23C]

| |
|---|
| [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. |

381. The combination of O'Neill and Jones discloses "*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*." Both O'Neill and Jones disclose initiating "*purchase transactions*."

382. This limitation of claim 23 is substantially the same limitation that is addressed in §V.M regarding claim limitation [12] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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. | [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. |

383. For the same reasons discussed above in §V.M for claim limitation [12], the combination of O'Neill and Jones discloses this limitation.

## V. Dependent Claim 30

### 1. Obtaining User Profile Information [30A]

| |
|---|
| [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 |

384. The combination of O'Neill and Jones discloses "*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*."

385. Jones discloses tracking user engagement with promotional content provided to the user, including items associated with the 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":

> For example, the database 18 may maintain a record of advertisements presented in response to queries containing particular keyword(s) and whether users interacted with the advertisements.
>
> The database 18 may also maintain information associated with a number of advertisements presented while voice queries are being processed by the system 10, a number of times a particular advertisement is presented in association with one or more keywords, advertisements hand-selected by providers, etc. For example, the server 16 may compute 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** (this information may be delivered back to the server 16 from the advertiser), etc.

(EX-1006, [0032]-[0033].)

386. O'Neill also discloses that the disclosed system 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:

> Preferably, the method includes a step of retrieving information from a previous communication stored in one or more of the data-storage units and combining the information with information from the current communication to enable the completion of a task associated with the previous communication.

(EX-1005, [0064].)

387. O'Neill further discloses:

> Furthermore, the improved dialogue manager provides a facility whereby the dialogue manager may retrieve information collected at an earlier point in a dialogue to assist in a current query.
>
> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user.

(EX-1005, [0076]-[0077].)

388. O'Neill describes the collection and use of information within a Knowledge Hierarchy:

> Referring to FIG. 2 b, the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session. In particular, the knowledge-storing components store data values elicited by the system from the user, or inferred by the system itself, in the course of a dialogue. These components can be collectively described as the Knowledge Hierarchy.
>
> The components in the Knowledge Hierarchy also maintain information that indicates how confirmed each datum is, and what the system intends to do (if anything) to confirm that datum adequately. The data values stored in the knowledge-storing elements are the system's discourse knowledge, and are used by components of the Expertise Hierarchy as they prepare system utterances in accordance with their generic and domain specific rules.

(EX-1005, [0103]-[0104].)

389. O'Neill also discloses:

> The above-described information-storage mechanism enables the improved dialogue manager **20** to store the dialogue product (i.e. the information elicited during the evolving dialogue of a dialogue session) in a tree-like structure (henceforth known as the knowledge tree). The tree-like structure of the dialogue product emerges from the fact that a dialogue frame may include as part of its Attributes, links to dialogue frames of other types (e.g. TheatreDialogFrame 133 may include a link to a PaymentDialogFrame 152).

(EX-1005, [0184].)

390. Obtaining and using user profile information to improve dialogue systems was well-known in the art as of the priority date, as evidenced by U.S. Patent 6,513,006 ("Howard"), which is discussed in greater detail below in §VIII. Howard discloses a "profile data store":

> The dialog manager works in conjunction with a profile data store 128. This data store may contain user profile information, such as data on what the user has previously requested. This information can be used, for example, to guide the dialog manager and the associated filtering function in identifying information the user may or may not be interested in. For example, if the user viewed a particular program on a previous night, that fact would be recorded in the data store 128 and the dialog manager could then suppress this program for display on a subsequent night. Similarly, if the user regularly rejects a dinner selection of spinach stew, the dialog manager will suppress this selection in the future. The profile data store can also be used to store user preferences about how much information the user wishes to have displayed on screen, or supplied by synthesized speech. The profile data for a particular user may be selected using speaker verification or speaker identification. In this regard, the recognition system stores unique or identifying speech characteristics of each user. Thereafter, when that user provides input speech, the system can either ascertain the identity of the speaker by performing speaker identification based on the speech characteristics, or verify a given identity provided by the speaker by performing speaker verification based on the speech characteristics.

(EX-1009, 6:31-54.)

391. Therefore, a POSA would be motivated to combinate the user tracking disclosed by O'Neill with Jones and create a system that also tracks the user's

interactions with promotional content, including items associated with the promotional content. The system disclosed by O'Neill tracks all user interactions with the system so it would be automatic for the O'Neill system to similarly track user interactions with promotional content.

392. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*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*." (§V.B.1.)

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

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

393. The combination of O'Neill and Jones discloses "*identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information*."

394. The '536 patent does not define the term "category" but does provide the following examples:

> For example, content/action identification module 235 may identify content of the voice-based input (e.g., words in the input), requested information (e.g., search results, a web page, music, video, graphics, or other information), requested actions (e.g., calculating a navigation route,

> placing a telephone call, playing a song, etc.), **a category or topic related to the input (e.g., music, business, stocks, sports, navigation, movies, etc.)**, or other criteria to use in selecting an advertisement.
> …
> For instance, a user may request airline reservations via voice-enabled device 210, and content/action identification module 235 may identify specific words used in the request, **a category related to the request (e.g., travel, airlines, hotels, etc.),** or other information.

(EX-1001, 6:17-40.)

395. Specifically, Jones discloses tracking user engagement with the promotional content provided to the user and using that to "identify[]…categories of items" to provide promotional content to the user:

> For example, the database 18 may maintain a record of advertisements presented in response to queries containing particular keyword(s) and whether users interacted with the advertisements.
>
> The database 18 may also maintain information associated with a number of advertisements presented while voice queries are being processed by the system 10, a number of times a particular advertisement is presented in association with one or more keywords, advertisements hand-selected by providers, etc. For example, the server 16 may compute 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 (this information may be delivered back to the server 16 from the advertiser), etc.

(EX-1006, [0032]-[0033].)

396. Therefore, the system disclosed by Jones track a user's interactions (e.g., "click throughs") with advertisements and stores that information in the database. That database is used when identifying what advertisements to provide to the user.

397. Similarly, Jones provides an example in which the user requests information related to restaurants (an activity as discussed above), and then identifies "*categories of items based on the prior user interactions specified by the user profile information*":

> **For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user.** The user telephone number or other 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.

(EX-1006, [0031].)

398. In this example, the Jones system is identifying the category of restaurants based on "*prior user interactions specified by the user profile information*."

399. O'Neill also discloses that the disclosed system 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:

> Preferably, the method includes a step of retrieving information from a previous communication stored in one or more of the data-storage units and combining the information with information from the current communication to enable the completion of a task associated with the previous communication.

(EX-1005, [0064].)

400. O'Neill further discloses:

> Furthermore, the improved dialogue manager provides a facility whereby the dialogue manager may retrieve information collected at an earlier point in a dialogue to assist in a current query.
>
> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user.

(EX-1005, [0076]-[0077].)

401. O'Neill describes the collection and use of information within a Knowledge Hierarchy:

> Referring to FIG. 2 b, the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session. In particular, the knowledge-storing components store data values elicited by the system from the user, or inferred by the system itself, in the course of a dialogue. These

> components can be collectively described as the Knowledge Hierarchy.
>
> The components in the Knowledge Hierarchy also maintain information that indicates how confirmed each datum is, and what the system intends to do (if anything) to confirm that datum adequately. The data values stored in the knowledge-storing elements are the system's discourse knowledge, and are used by components of the Expertise Hierarchy as they prepare system utterances in accordance with their generic and domain specific rules.

(EX-1005, [0103]-[0104].)

402. O'Neill also discloses:

> The above-described information-storage mechanism enables the improved dialogue manager **20** to store the dialogue product (i.e. the information elicited during the evolving dialogue of a dialogue session) in a tree-like structure (henceforth known as the knowledge tree). The tree-like structure of the dialogue product emerges from the fact that a dialogue frame may include as part of its Attributes, links to dialogue frames of other types (e.g. TheatreDialogFrame 133 may include a link to a PaymentDialogFrame 152).

(EX-1005, [0184].)

403. Therefore, a POSA would be motivated to combinate the user tracking disclosed by O'Neill with Jones and create a system that also tracks the user's interactions with promotional content. The system disclosed by O'Neill tracks all user interactions with the system so it would be automatic for the O'Neill system to similarly track user interactions with promotional content.

404. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information*." (§V.B.1.)

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

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

405. The combination of O'Neill and Jones discloses "*wherein determining the promotional content comprises determining a promotional item associated with one of the categories*."

406. As discussed above regarding limitation [30B], Jones discloses "*identifying categories of items based on prior user interactions specified by the user profile information*." Also as discussed above regarding limitation [30B], the combination of O'Neill and Jones would use that information to determine and provide the user with promotional content associated with one or more of those categories.

407. Jones discloses a system that determines "*determining a promotional item associated with one of the categories*," because the disclosed system uses a number of factors to weight and determine the appropriate "*promotional item*," including "*categories*":

> An advertisement presented to a user may be based on a weight associated with the advertisement in accordance with a single or combined factors such as advertiser contract commitments, bidding price of advertisers, popularity with users, keyword mapping to advertisements, statistical usage (e.g. least recently presented), user demographics, provider choice of advertisement, geographical location of the user and/or provider, etc. For example, a voice query pertaining to a schedule of performances at The Kennedy Center for the Performing Arts in Washington, D.C. from a user who indicates being a resident of New York or whose telephone number corresponds to a New York number, may trigger one or more advertisements related to Broadway shows in New York to be presented to the user, who might respond by purchasing tickets.

(EX-1006, [0029].)

408. In this example, the user is interested in the activity of attending a performance at The Kennedy Center for the Performing Arts and the system determined promotional activity—advertisements related to Broadway shows—based on that activity. (EX-1006, [0029].)

409. Similarly, Jones provides an example in which the user requests information related to restaurants (an activity as discussed above), and then "*determin[es] the promotional content based on the activity*":

> **For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user.** The user telephone number or other 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.

(EX-1006, [0031].)

410. A POSITA would also use this teaching from Jones in combination with the example provided in O'Neill above. (§V.C.2.) In that example, the user requested assistance booking a cinema ticket. (EX-1005, [0391]-[0430].) A POSITA would use the teachings of Jones to determine promotional content based on that activity and provide the user promotional content related to cinema tickets.

## W. Independent Claim 50

### 1. Preamble [50P]

| |
|---|
| [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: |

411. The combination of O'Neill and Jones discloses "*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*."

412. This preamble of claim 50 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 1 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

413. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill and Jones discloses this limitation.

### 2. Receiving a First Natural Language Utterance [50A]

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

414. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a first natural language utterance*."

415. This limitation of claim 50 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [50A] receiving, at the one or more physical processors, a first natural language utterance; |

416. For the same reasons discussed above in §V.B.2 for claim limitation [1A], the combination of O'Neill and Jones discloses this limitation.

### 3. Providing a Response to the First Natural Language Utterance [50B]

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

417. The combination of O'Neill and Jones discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

418. This limitation of claim 50 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [50B] providing, by the one or more physical processors, a response to the first natural language utterance; |

419. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill and Jones discloses this limitation.

### 4. Receiving a Second Natural Language Utterance [50C]

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

420. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*."

421. This limitation of claim 50 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

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

422. For the same reasons discussed above in §V.B.4 for claim limitation [1C], the combination of O'Neill and Jones discloses this limitation.

### 5. Obtaining User Profile Information [50D]

| |
|---|
| [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; |

423. The combination of O'Neill and Jones discloses "*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*."

424. This limitation of claim 50 is substantially the same limitation that is addressed in §V.V.1 regarding claim limitation [30A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

425. For the same reasons discussed above in §V.V.1 for claim limitation [30A], the combination of O'Neill and Jones discloses this limitation.

### 6. Identifying Categories of Items Based on Prior User Interactions [50E]

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

426. The combination of O'Neill and Jones discloses "*identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information*."

427. This limitation of claim 50 is substantially the same limitation that is addressed in §V.V.2 regarding claim limitation [30B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [30B] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information, | [50E] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information; |

428. For the same reasons discussed above in §V.V.2 for claim limitation [30B,] the combination of O'Neill and Jones discloses this limitation.

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

| |
|---|
| [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 |

429. The combination of O'Neill and Jones discloses "*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*."

430. O'Neill describes that the system engages in dialogue with the user in which past and present utterances from the user and system are processed to formulate a response, including "*the first natural language utterance and the second natural language utterance*":

> Since the dialogue frame taken from the DiscourseStack 123 details the information previously provided by the user together with the status of the information (repeated, modified, negated, etc.) and the system's previous intentions for confirming or repairing supplied or missing information, the generic confirmation strategy enables the improved dialogue manager 20 to interpret the user's most recent utterance in light of the improved dialogue manager's own last utterance. For instance, if the user answers "Yes" in response to the system's observation "So, you're staying in Dublin", then Dublin can be regarded as the confirmed location.
>
> As will be recalled, the domain-specific 'Experts' all inherit the generic dialogue handling strategies of the DiscourseManager 22. In particular, in order for each Expert to update the record of the evolving dialogue, each Expert takes on the improved dialogue manager's 120 DiscourseHistory 122 as an inheritable attribute of its own and the evolving domain-specific and generally Expert-specific frames of attributes are maintained on the DiscourseStack 123 within the DiscourseHistory object 122.

(EX-1005, [0213]-[0214].)

431. Further, as discussed in §V.C above, O'Neill describes "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-1005, [0103]-[0104].)

432. As discussed above regarding limitation [1K], Jones discloses "*determining, by the one or more physical processors, promotional content based on the interpretation*" of the second natural language utterance. (*Supra* §V.B.11.) Also as discussed above regarding limitation [1L], the combination of O'Neill and Jones would use that information to determine and provide the user with promotional content associated with one or more requests. (*Supra* §V.B.12.) Jones would similarly determine promotional content "*based on the first natural language utterance and the second natural language utterance*."

433. Jones discloses using keywords from the user's requests to target advertisement: "The system and method disclosed 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].)

434. Jones discloses a system that determines promotional content "*based on the first natural language utterance and the second natural language utterance*,"

because the disclosed system uses a number of factors to weigh and determine the appropriate "*promotional content*," including the user's requests and other factors:

> An advertisement presented to a user may be based on a weight associated with the advertisement in accordance with a single or combined factors such as advertiser contract commitments, bidding price of advertisers, popularity with users, keyword mapping to advertisements, statistical usage (e.g. least recently presented), user demographics, provider choice of advertisement, geographical location of the user and/or provider, etc. For example, a voice query pertaining to a schedule of performances at The Kennedy Center for the Performing Arts in Washington, D.C. from a user who indicates being a resident of New York or whose telephone number corresponds to a New York number, may trigger one or more advertisements related to Broadway shows in New York to be presented to the user, who might respond by purchasing tickets.

(EX-1006, [0029].)

> For example, the database 18 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].)

435. A POSITA would use this teaching from Jones in combination with the example provided in O'Neill above. (§V.A.1.) A POSITA would use the teachings of O'Neill that the system should use past and present user utterances to process a response with the teachings of Jones to determine promotional content based on relevant information about the user's utterance and other factors. Therefore, the

combined system would "*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*."

### 8. Presenting Promotional Content [50G]

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

436. The combination of O'Neill and Jones discloses "*presenting, by the one or more physical processors, the promotional content to the user*."

437. This limitation of claim 50 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [50G] presenting, by the one or more physical processors, the promotional content to the user. |

438. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill and Jones discloses this limitation.

## X. Independent Claim 53

### 1. Preamble [53P]

| |
|---|
| [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: |

439. To the extent the preamble is limiting, the combination of O'Neill and Jones discloses "*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*."

440. This preamble of claims 49 and 53 are substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 1 above. It is also the same preamble that appears in claims 43 and 49. I provide a side-by-side comparison of the limitations here:

| | | |
|---|---|---|
| [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: | [43P] / [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: | [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: |

441. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill and Jones discloses this limitation.

### 2. Receive a First Natural Language Utterance [53A]

| |
|---|
| [53A] receive a first natural language utterance; |

442. The combination of O'Neill and Jones discloses "*receive a first natural language utterance*."

443. This limitation of claim 53 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [53A] receive a first natural language utterance; |

444. For the same reasons discussed above in §V.B.2 for claim limitation [1A], the combination of O'Neill and Jones discloses this limitation.

### 3. Provide Response to First Natural Language Utterance [53B]

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

445. The combination of O'Neill and Jones discloses "*provide a response to the first natural language utterance*."

446. This limitation of claim 53 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [53B] provide a response to the first natural language utterance; |

447. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill and Jones discloses this limitation.

### 4. Receive a Second Natural Language Utterance [53C]

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

448. The combination of O'Neill and Jones discloses "*receive a second natural language utterance relating to the first natural language utterance*."

449. This limitation of claim 53 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance; | [53C] receive a second natural language utterance relating to the first natural language utterance; |

450. For the same reasons discussed above in § V.B.4 for claim limitation [1C], the combination of O'Neill and Jones discloses this limitation.

### 5. Obtaining User Profile Information [53D]

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

451. The combination of O'Neill and Jones discloses "*obtain user profile information associated with a user, wherein the user profile information specifies prior user interactions with items*."

452. This limitation of claim 53 is substantially the same limitation that is addressed in §V.V.1 regarding claim limitation [30A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

453. For the same reasons discussed above in § V.V.1 for claim limitation [30A], the combination of O'Neill and Jones discloses this limitation.

### 6. Identify Categories of Items Based on Prior User Interactions [53E]

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

454. The combination of O'Neill and Jones discloses "*identify categories of items based on the prior user interactions specified by the user profile information*."

455. This limitation of claim 53 is substantially the same limitation that is addressed in §V.V.2 regarding claim limitation [30B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [30B] identifying, by the one or more physical processors, categories of items based on the prior user interactions specified by the user profile information, | [53E] identify categories of items based on the prior user interactions specified by the user profile information; |

456. For the same reasons discussed above in §V.V.2 for claim limitation [30B], the combination of O'Neill and Jones discloses this limitation.

### 7. Determine Promotional Content Associated with One of the Categories [53F]

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

457. The combination of O'Neill and Jones discloses "*determine, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories*."

458. This limitation of claim 53 is substantially the same limitation that is addressed in §V.V.3 regarding claim limitation [30C] above. It is also substantially the same as limitation [50F] discussed above. I provide a side-by-side comparison of the limitations here:

| | | |
|---|---|---|
| [30C] wherein determining the promotional content comprises determining a promotional item associated with one of the categories. | [53F] determine, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and | [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 |

459. For the same reasons discussed above in §V.V.3 for claim limitation [30C] and §V.W.7 for claim limitation [50F], the combination of O'Neill and Jones discloses this limitation.

### 8. Present Promotional Content [53G]

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

460. The combination of O'Neill and Jones discloses "*present the promotional content to the user*."

461. This limitation of claim 53 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [49G] present the promotional content to a user. |

462. For the same reasons discussed above in § V.B.12 for claim limitation [1L], the combination of O'Neill and Jones discloses this limitation.

## Y. Independent Claim 54

### 1. Preamble [54P]

| |
|---|
| [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: |

463. The combination of O'Neill and Jones discloses "*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*."

464. This preamble of claim 54 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 1 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

465. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill and Jones discloses this limitation.

### 2. Receiving a First Natural Language Utterance [54A]

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

466. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a first natural language utterance*."

467. This limitation of claim 54 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [54A] receiving, at the one or more physical processors, a first natural language utterance; |

468. For the same reasons discussed above in §V.B.2 for claim limitation [1A], the combination of O'Neill and Jones discloses this limitation.

### 3. Providing a Response to the First Natural Language Utterance [54B]

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

469. The combination of O'Neill and Jones discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

470. This limitation of claim 54 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [54B] providing, by the one or more physical processors, a response to the first natural language utterance; |

471. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill and Jones discloses this limitation.

### 4. Receiving a Second Natural Language Utterance [54C]

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

472. The combination of O'Neill and Jones discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*."

473. This limitation of claim 54 is substantially the same limitation that is addressed in §V.B.5 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

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

474. For the same reasons discussed above in §V.B.5 for claim limitation [1C], the combination of O'Neill and Jones discloses this limitation.

### 5. Obtaining User Profile Information [54D]

| |
|---|
| [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; |

475. The combination of O'Neill and Jones discloses "*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*."

476. This limitation of claim 54 is substantially the same limitation that is addressed in §V.V.1 regarding claim limitation [30A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

477. For the same reasons discussed above in §V.V.1 for claim limitation [30A], the combination of O'Neill and Jones discloses this limitation.

### 6. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [54E]

| |
|---|
| [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; |

478. The combination of O'Neill and Jones discloses "*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*."

479. This limitation of claim 54 is substantially the same limitation that is addressed in §V.N.2 regarding claim limitation [16B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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. | [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; |

480. For the same reasons discussed above in §V.N.2 for claim limitation [30A], the combination of O'Neill and Jones discloses this limitation. As discussed

above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*identifying . . . one or more requests associated with the first natural language utterance or the second natural language utterance*." (*See supra* §V.B.1.)

### 7. Determining Application for Processing One or More Requests [54F]

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

481. The combination of O'Neill and Jones discloses "*determining, by the one or more physical processors, one or more applications for processing the one or more requests*."

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 "AccomodationExpert," 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 AccomodationExpert 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].)

### 8. Identifying Categories of Items Based on Prior Use Interactions Related to One or More Applications [54G]

| |
|---|
| [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; |

483. The combination of O'Neill and Jones discloses "*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*."

484. This limitation of claim 54 is substantially the same limitation that is addressed in §V.W.7 regarding claim limitation [50F] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

485. For the same reasons discussed above in §V.W.7 for claim limitation [31D], the combination of O'Neill and Jones discloses this limitation.

### 9. Determining Promotional Content Associated with One of the Categories [54H]

| |
|---|
| [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 |

486. The combination of O'Neill and Jones discloses "*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*."

487. This limitation of claim 54 is substantially the same limitation that is addressed in §V.V.3 regarding claim limitation [50F] above. II provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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: |

488. For the same reasons discussed above in §V.W.7 for claim limitation [50F], the combination of O'Neill and Jones discloses this limitation.

### 10. Presenting the Promotional Content [54I]

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

489. The combination of O'Neill and Jones discloses "*presenting, by the one or more physical processors, the promotional content to the user*."

490. This limitation of claim 54 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [54I] presenting, by the one or more physical processors, the promotional content to the user. |

491. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill and Jones discloses this limitation.

## Z. Independent Claim 55

### 1. Preamble [55P]

| |
|---|
| [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: |

492. To the extent the preamble is limiting, the combination of O'Neill and Jones discloses "*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*."

493. This preamble of claim 55 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 1 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

494. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill and Jones discloses this limitation.

### 2. Receive a First Natural Language Utterance [55A]

| |
|---|
| [55A] receive a first natural language utterance; |

495. The combination of O'Neill and Jones discloses "*receive a first natural language utterance*."

496. This limitation of claim 55 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [53A] receive a first natural language utterance; |

497. For the same reasons discussed above in §V.B.2 for claim limitation [1A], the combination of O'Neill and Jones discloses this limitation.

### 3. Provide Response to First Natural Language Utterance [55B]

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

498. The combination of O'Neill and Jones discloses "*provide a response to the first natural language utterance*."

499. This limitation of claim 55 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [55B] provide a response to the first natural language utterance; |

500. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill and Jones discloses this limitation.

### 4. Receive a Second Natural Language Utterance [55C]

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

501. The combination of O'Neill and Jones discloses "*receive a second natural language utterance relating to the first natural language utterance*."

502. This limitation of claim 55 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance; | [55C] receive a second natural language utterance relating to the first natural language utterance; |

503. For the same reasons discussed above in §V.B.4 for claim limitation [1C], the combination of O'Neill and Jones discloses this limitation.

### 5. Obtain User Profile Information [55D]

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

504. The combination of O'Neill and Jones discloses "*obtain user profile information associated with a user, wherein the user profile information specifies prior user interactions with items*."

505. This limitation of claim 55 is substantially the same limitation that is addressed in §V.V.1 regarding claim limitation [30A] above. I provide a side-by-side comparison of the limitations here:

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

506. For the same reasons discussed above in §V.V.1 for claim limitation [30A], the combination of O'Neill and Jones discloses this limitation.

### 6. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [55E]

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

507. The combination of O'Neill and Jones discloses "*identify one or more requests associated with one or more of the first natural language utterance or the second natural language utterance*."

508. This limitation of claim 31 is substantially the same limitation that is addressed in §V.N.2 regarding claim limitations [16B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [16B] identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance, | [55E] identify one or more requests associated with one or more of the first natural language utterance or the second natural language utterance; |

509. While claim limitation [16B] requires identifying "*one or more requests associated with the second natural language utterance*," claim limitation [55E] (and [31B]) requires identifying "one or more requests associated with the first natural language utterance ***or*** the second natural language utterance." Therefore, for the same reasons discussed above in §V.N.2 for claim limitation [16B], the combination of O'Neill and Jones discloses this limitation.

510. Moreover, O'Neill also discloses identifying requests associated with the first natural language utterance. For example, O'Neill discloses an example

where a user's first natural language utterance is "I'd like to go see THE LION KING 2," and in response the system identifies that the user is making a request for assistance booking a cinema ticket and use the CinemaExpert domain agent to process the user's request. (EX-1005, [0203]-[0204].)

### 7. Determine Application for Processing One or More Requests [55F]

| |
|---|
| [55E] determine one or more applications for processing the one or more requests; and |

511. The combination of O'Neill and Jones discloses "*determine one or more applications for processing the one or more requests*."

512. This limitation of claim 55 is substantially the same limitation that is addressed in §V.Y.7 regarding claim limitation [54F] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [54F] determining, by the one or more physical processors, one or more applications for processing the one or more requests; and | [55F] determine one or more applications for processing the one or more requests; and |

513. For the same reasons discussed above in §V.Y.7 for claim limitation [54F], the combination of O'Neill and Jones discloses this limitation.

### 8. Identify Categories of Items Based on Prior Use Interactions Related to One or More Applications [55G]

| |
|---|
| [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; |

514. The combination of O'Neill and Jones discloses "*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*."

515. This limitation of claim 55 is substantially the same limitation that is addressed in §V.Y.8 regarding claim limitation [55G] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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; | [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; |

516. For the same reasons discussed above in §V.Y.8 for claim limitation [54G], the combination of O'Neill and Jones discloses this limitation.

### 9. Determine Promotional Content Associated with One of the Categories [55H]

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

517. The combination of O'Neill and Jones discloses "*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*."

518. This limitation of claim 55 is substantially the same limitation that is addressed in §V.V.3 regarding claim limitation [30C] above. It is also substantially the same as limitation [50F], [53F] and [54H] discussed above. I provide a side-by-side comparison of the limitations here:

| | | |
|---|---|---|
| [30C] wherein determining the promotional content comprises determining a promotional item associated with one of the categories. | [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 | [55H] determine, based on the first natural language utterance and the second natural language utterance, promotional content associated with one of the categories; and |

519. For the same reasons discussed above in §V.V.3 for claim limitation [30C] and §V.W.7 for claim limitation [50F], the combination of O'Neill and Jones discloses this limitation.

### 10. Present Promotional Content [55I]

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

520. The combination of O'Neill and Jones discloses "*present the promotional content to the user*."

521. This limitation of claim 55 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [49G] present the promotional content to a user. |

522. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill and Jones discloses this limitation.

## VI. GROUND 2: THE COMBINATION OF O'NEILL, JONES AND FRANCO RENDERS OBVIOUS CLAIMS 13-15, 24 AND 33-38

### A. Overview of the Combination

523. The combination of O'Neill, Jones and U.S. Patent Application Publication No. US 2010/0125458 ("Franco"; EX-1007) discloses every limitation of claims 13-15, 24 and 33-38. O'Neill, Jones and Franco are analogous art to the '536 patent because all three references are directed to speech recognition systems and software.

524. I have described both O'Neill and Jones above, and I discuss Franco below.

#### 1. Franco

525. Franco "relates generally to speech recognition applications and relates more particularly to correcting errors in speech recognition systems." (EX-1007, [0002].) Franco discloses a "method and apparatus for error correction in speech recognition applications" that "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].) "It is assumed that, if the first utterance has been incorrectly recognized on a first attempt, the user will repeat

the first utterance (or at least the incorrectly recognized portion of the first utterance) in the subsequent utterance." (EX-1007, [0006].) "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." (EX-1007, [0020].)

526. Franco discloses at least three ways of evaluating a combined hypothetical analysis of a corrected utterance:

> [I]f the method 100 determines in step 110 that a need for correction has been indicated, the method 100 proceeds to step 112 and produces a subsequent hypothesis based on a subsequent utterance (e.g., a "correction" utterance). In one embodiment, the subsequent hypothesis is produced in accordance with one or more known speech recognition techniques. The method 100 then proceeds to step 114 and combines the first hypothesis with the subsequent hypothesis to form a combined hypothesis. In one embodiment, the method 100 specifically combines the respective acoustic scores of the first hypothesis and the subsequent hypothesis. In one embodiment, the respective acoustic scores are combined at the frame level (e.g., using a weighted combination of log observation probabilities). In another embodiment, the respective acoustic scores are combined at the re-scoring level (e.g., where either N-best or lattices from independent recognition attempts for each hypothesis are combined).

(EX-1007, [0017].)

### 2. Motivation to Combine

527. A POSITA would have been motivated to integrate the error correction features speech recognition disclosed in Franco with the speech recognition systems disclosed in O'Neill and Jones for a number of reasons.

528. First, O'Neill recognizes the motivation to combine the disclosed system with other dialogue systems like Franco: "It will be appreciated 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 or other transactional systems." (EX-1005, [0089].) A POSITA would understand O'Neill expressing a motivation to combine the system disclosed in O'Neill with additional systems, like Franco, that specialize in correcting errors in speech recognition systems. O'Neill also recognizes that its system builds on the DARPA Communicator system architecture but O'Neill does not provide all of the details of the operation of that DARPA Communicator 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.

529. 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 the method of error correction disclosed in Franco was broadly applicable and could be implemented and combined with speech recognition applications, like O'Neill and Jones.

530. Third, a POSITA would be motivated to reduce the number of errors in a speech recognition system. A POSITA designing a speech recognition system would seek out systems that could be useful in detecting and correcting errors in speech recognition system. Therefore, it would be logical for a POSITA to look to a reference like Franco when designing a speech recognition system like those described in O'Neill and Jones.

531. Fourth, the results of the combination would have been predictable to a POSITA and the 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].) The systems disclosed in O'Neill and Jones are implemented through the a communication channel, like a telephone system.

## B. Dependent Claim 13

### 1. Interactive Session [13A]

| |
|---|
| [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: |

532. The combination of O'Neill, Jones and Franco discloses "*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*."

533. This limitation of claim 13 is substantially the same limitation that is addressed in §V.C.1 regarding claim limitation [2A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

534. For the same reasons discussed above in §V.C.1 for claim limitation [2A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Receiving a Third Natural Language Utterance During the Interactive Session [13B]

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

535. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a third natural language utterance during the interactive session*."

536. This limitation of claim 13 is substantially the same limitation that is addressed in §V.I regarding claim limitation [8] above. I provide a side-by-side comparison of the limitations here:

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

537. For the same reasons discussed above in §V.I for claim [8], the combination of O'Neill, Jones and Franco discloses this limitation.

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

| |
|---|
| [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. |

538. The combination of O'Neill, Jones and Franco discloses "*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*."

539. O'Neill discloses "*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*," as O'Neill

discloses a system that includes what it 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:

> The DiscourseManager 22 implements the generic confirmation strategy and creates new dialogue frames for each utterance of the user by determining whether the utterance represents the repetition, modification or negation of information previously provided by the user. The DiscourseManager 22 performs this determination by comparing the dialogue frame corresponding to the user's latest utterance, with a corresponding dialogue frame representing the last discourse state taken from the DiscourseStack 123.
>
> Since the dialogue frame taken from the DiscourseStack 123 details the information previously provided by the user together with the status of the information (repeated, modified, negated, etc.) and the system's previous intentions for confirming or repairing supplied or missing information, the generic confirmation strategy enables the improved dialogue manager 20 to interpret the user's most recent utterance in light of the improved dialogue manager's own last utterance. For instance, if the user answers "Yes" in response to the system's observation "So, you're staying in Dublin", then Dublin can be regarded as the confirmed location.

(EX-1005, [0212]-[0213].)

540. Therefore, O'Neill discloses processing 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.

541. O'Neill further discloses the use of confirmation strategy to identify and reduce the risk of "misinterpretations of users' free-form utterances" (EX-1005, [0236]):

> Before transferring (or refusing a transfer) to a new handling agent, the improved dialogue manager always confirms the user's intentions ("Do you want to enquire about cinema bookings?"). Bearing in mind that the improved dialogue manager can operate in a number of different though potentially related transaction domains, the above confirmation strategy reduces the risk of serious misinterpretations of users' free-form utterances.

(EX-1005, [0236].)

542. Franco likewise discloses "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."

543. Franco discloses:

> In one embodiment, the present invention is a method and apparatus for error correction in speech recognition applications. In one embodiment, 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. It is assumed that, if the first utterance has been incorrectly recognized on a first attempt, the user will repeat the first

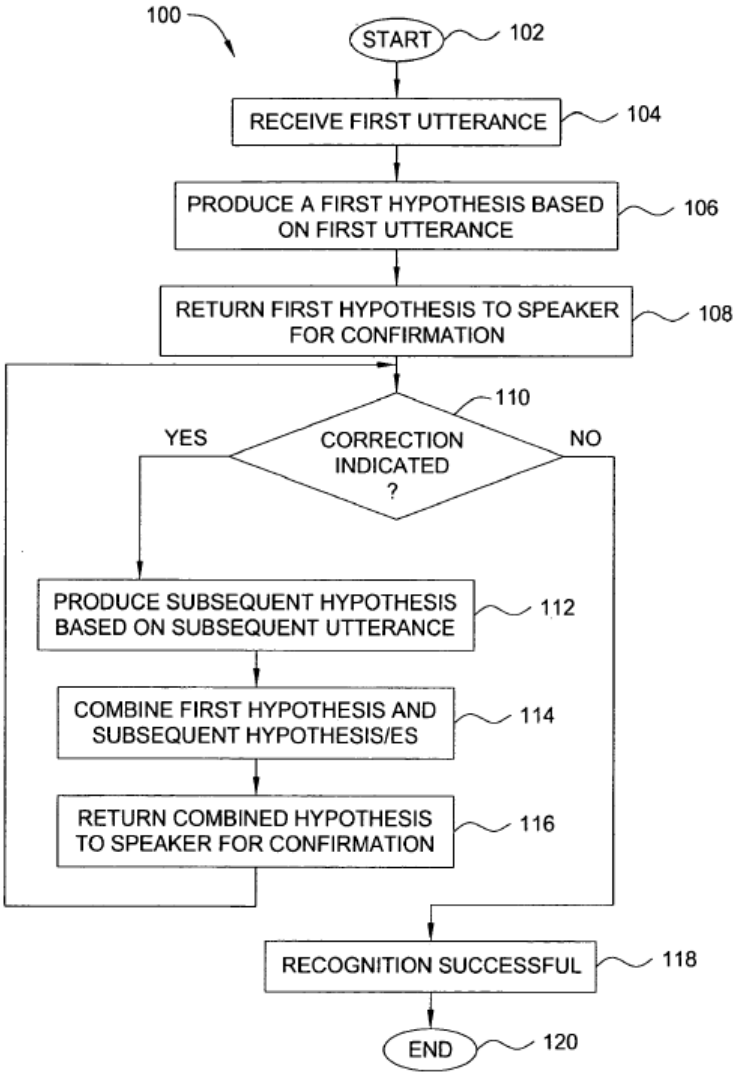
> utterance (or at least the incorrectly recognized portion of the first utterance) in the subsequent utterance.

(EX-1007, [0006].)

544. Franco "relates generally to speech recognition applications and relates more particularly to correcting errors in speech recognition systems." (EX-1007, [0002].) Errors can be found when, for example, "background noise present at the time the utterance is made or speaker enunciation [] affect[s] a speech recognition application's ability to produce accurate results in a timely fashion." (EX-1007, [0004].)

545. The system in Franco assumes that "when speech recognition is unsuccessful on the first attempt, a speaker will naturally repeat all or part of the same utterance in order to correct the recognition error. Thus, acoustic evidence (e.g., recognition hypotheses) for multiple instances of the same utterance may be combined in a manner that produces a better estimate of the utterance." (EX-1007, [0012].)

546. For example, Franco's Figure 1 provides a "flow diagram illustrating one embodiment of a method for correcting errors in a speech recognition process." (EX-1007, [0007].) In Figure 1, below, the system receives an utterance (104) and provides a hypothesized response to the speaker (108).

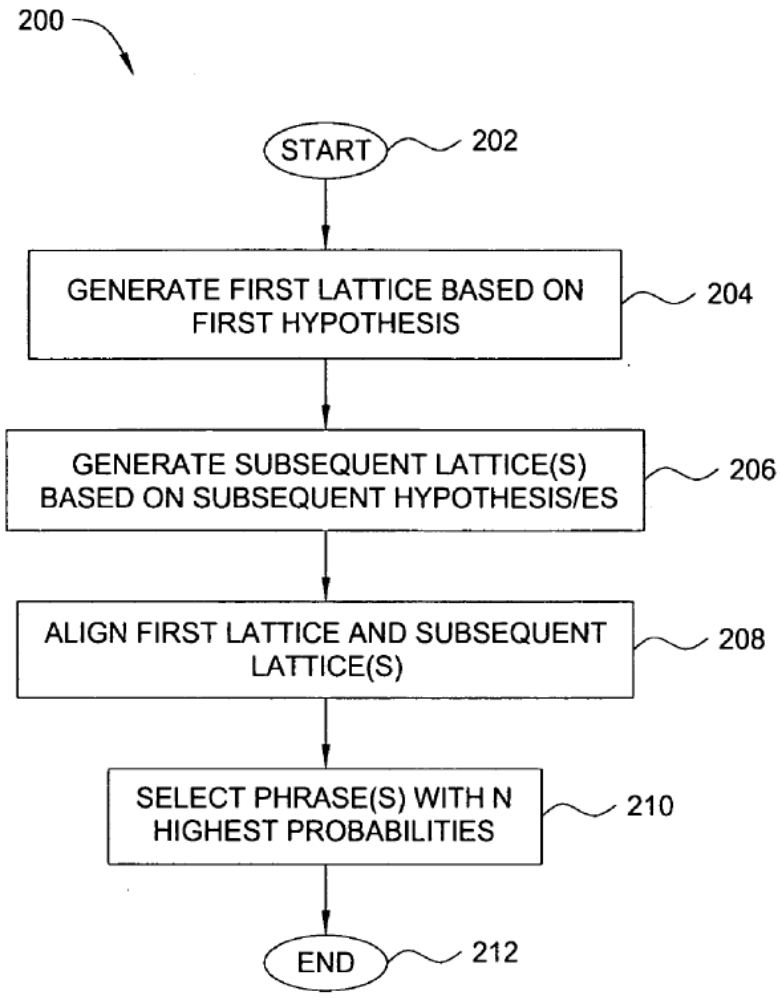



**Franco, Figure 1**

547. Franco explains that, if the system "determines in step 110 that a need for correction has been indicated, the method 100 proceeds to step 112 and produces a subsequent hypothesis based on a subsequent utterance (e.g., a 'correction utterance')." (EX-1007, [0017].) "[T]he combined hypothesis may be produced on the basis of a single utterance that comprises a repeated phrase (including a keyword that indicates the repetition)." Franco explains that, "[i]n such an instance, the

speaker essentially provides the 'subsequent' utterance without an intervening prompt from the method 100" and therefore "there is no intervening prompt from the method 100 (e.g., such as providing the first hypothesis to the speaker)." (EX-1007, [0018].) In such case, "the method 100 returns the combined hypothesis to the speaker for review/confirmation." (EX-1007, [0019].) The method then "determine[s] whether further correction to the combined hypothesis is necessary," and "continues to combine hypotheses in an effort to correct recognition of the first utterance until recognition is determined to be successful (e.g., step 118)." (EX-1007, [0019].)

548. Franco's Figure 2, below, is "a flow diagram illustrating one embodiment of a method for combining hypotheses, according to the present invention." (EX-1007, [0009].)




**Franco, Figure 2**

549. Franco explains that "the method 200 selects the phrase(s) with the N highest probabilities as the combined hypothesis/es to be presented to the speaker (e.g., in accordance with step 116 of the method 100)." (EX-1007, [0026].) "In another embodiment, if the phrase with the highest probability is identical to the first hypothesis, which the speaker has indicated as being incorrect, then the method 200 selects the phrase with the second highest probability as an alternative combined hypothesis." (EX-1007, [0026].)

550. In this example, the system reinterprets the misrecognized words from the first utterance in response to subsequent utterances. (EX-1007, [0020].) 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, Franco discloses an "*adaptive misrecognition engine*."

551. Franco also discloses:

> The method 200 is initialized at step 202 and proceeds to step 204, where the method 200 generates a first word lattice based on the first hypothesis. The first word lattice is a network or diagram that graphically illustrates the first hypothesis. Specifically, the first lattice may be thought of as comprising a plurality of lattice positions or word "slots", where each slot corresponds to a temporal interval in the first utterance. Each slot is associated with one or more words that may potentially be have been uttered during the corresponding temporal interval. All of the slots in total comprise a phrase that may represent the first utterance (e.g., by selecting one word for each slot and then concatenating the selected words). Each potential word is associated with a score that indicates the probability that the word was in fact, uttered during the temporal interval corresponding to that slot. In one embodiment, a deletion may be the most probable word at one or more slots in the first word lattice. In fact, if the waveform representing the first utterance has a low signal-to-noise ratio, the first word lattice may include a plurality of slots whose most probable word is a deletion, due to multiple insertions with low probabilities.

(EX-1007, [0022].)

552. While Franco does not explicitly "determin[ing] a reinterpretation of the second natural language utterance," the Franco system does disclose using a

subsequent utterance (a "third natural language utterance") to "determine a reinterpretation of" a prior natural language utterance ("the second natural language utterance").

553. Franco discloses a process for reinterpreting misrecognized utterances that mirrors the support for the alleged invention in the '536 patent. The '536 patent explains that the "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, entitled 'System and Method of Supporting Adaptive Misrecognition in Conversational Speech,' … which issued as U.S. Pat. No. 7,620,549" ["Di Cristo"]. (EX-1001, 5:14-25-64.)

554. Di Cristo explains that "the parser is further configured to generate a plurality of interpretations of the natural language utterance," and "the analyze is further configured to use the personalized cognitive model to select a next best interpretation from the plurality of interpretations **if the adaptive misrecognition engine determines that the interpretation generated by the parser is incorrect**." (EX-1012, 33:42-48.) Therefore, Franco discloses a process for reinterpreting misrecognized utterances that mirrors the support for the alleged invention in the '536 patent.

### C. Dependent Claim 14

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

555. The combination of O'Neill, Jones and Franco discloses "*wherein determining the promotional content comprises determining the promotional content based on the reinterpretation*."

556. As discussed above regarding limitation [1K] (§V.B.11), Jones discloses "*determining, by the one or more physical processors, promotional content based on the interpretation*" of the second natural language utterance. Also as discussed above regarding limitation [1L] (§V.B.12), the combination of O'Neill and Jones would use that information to determine and provide the user with promotional content associated with one or more requests. To the extent reinterpretation of the second language utterance was necessary, the system would use that reinterpretation to select an advertisement.

557. Jones discloses using keywords from the user's request to target advertisement: "The system and method disclosed 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].)

558. Jones discloses a system that determines "*promotional content based on the reinterpretation*," because the disclosed system uses a number of factors to weight and determine the appropriate "*promotional content*," including the user's requests and other factors:

> An advertisement presented to a user may be based on a weight associated with the advertisement in accordance with a single or combined factors such as advertiser contract commitments, bidding price of advertisers, popularity with users, keyword mapping to advertisements, statistical usage (e.g. least recently presented), user demographics, provider choice of advertisement, geographical location of the user and/or provider, etc. For example, a voice query pertaining to a schedule of performances at The Kennedy Center for the Performing Arts in Washington, D.C. from a user who indicates being a resident of New York or whose telephone number corresponds to a New York number, may trigger one or more advertisements related to Broadway shows in New York to be presented to the user, who might respond by purchasing tickets.

(EX-1006, [0029].)

559. In this example, the user has made a request for information regarding attending a performance at The Kennedy Center for the Performing Arts and the system determined promotional content—advertisements related to Broadway shows—based on that request. (EX-1006, [0029].)

560. Similarly, Jones provides an example in which the user requests information related to restaurants (an activity as discussed above), and then

"*determin[es] promotional content that relates to one or more of the first request or the second request*":

> **For example, when a user requests information pertaining to "restaurants in Indiana" and data records in the database 18 indicates that a particular seafood restaurant advertisement has been presented to similar previous queries, the seafood restaurant advertisement may be presented to the user.** The user telephone number or other 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.

(EX-1006, [0031].)

561. A POSITA would also use this teaching from Jones in combination with the example provided in O'Neill and Franco above. (§V.B.) A POSITA would use the teachings of Franco to reinterpret a natural language utterance and use the reinterpretation to provide the user a more targeted advertisement.

## D. Dependent Claim 15

### 1. Determining the Promotional Content Based on the Interpretation [15A]

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

562. The combination of O'Neill, Jones and Franco discloses "*wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising*."

563. Claim 15 depends from claim 13, which depends from claim 1. The "determining the promotional content based on the interpretation" is described above for claim 1 in §V.B.11. This claim limitation is substantially the same limitation that is addressed in §V.B.11 regarding claim limitation [1K] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1K] determining, by the one or more physical processors, promotional content based on the interpretation: | [15A] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising |

564. For the same reasons discussed above in §V.B.11 for claim limitation [1K], the combination of O'Neill, Jones and Franco discloses this limitation.

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

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

565. The combination of O'Neill, Jones and Franco discloses "*determining, by the one or more physical processors, other promotional content based on the reinterpretation*."

566. As discussed above (§V.L), the combination of O'Neill, Jones and Franco will use the reinterpreted utterance to determine what promotional content to provide to the user.

567. Jones provides examples where "*other promotional content*" is determined and presented "*based on [a] reinterpretation*" of the user's utterance:

> After transcribing 54 keyword(s), process 50 continues to selecting 56 advertisement data for the transcribed keyword(s). For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes.

(EX-1006, [0040].)

568. Jones recognizes that some words, like "apple" may have multiple meaning or connotations to a user and qualification may be necessary:

> The database 18 (FIG. 1) may maintain a histogram of usage for multiple meanings and/or forms of keyword(s) or phrase(s) which may be used to prompt a user (or consumer) or provider to qualify or clarify keyword(s) or phrase(s) having multiple meanings or forms when formulating a search query and an advertisement may be selected based on the qualified meaning of the keyword(s) or phrase(s).

(EX-1006, [0053].)

569. Jones provides an example where other promotional content may be determined and presented based on the reinterpretation or qualification:

> For example, when a search query containing the keyword "apple" is qualified to mean the fruit, an advertisement pertaining to food recipe and/or the fruit may be presented, while an advertisement pertaining to the tax program and/or tax tips are provided when the keyword "apple" in the query is qualified to mean the tax program.

(EX-1006, [0058].)

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

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

570. The combination of O'Neill, Jones and Franco discloses "*presenting, by the one or more physical processors, the other promotional content to the user*."

571. As discussed above in §V.D.2 related to claim limitation [15B], once the system determines "other promotional content based on the reinterpretation," that promotional content is then presented to the user.

## E. Dependent Claim 24

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

| |
|---|
| [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; |

572. The combination of O'Neill, Jones and Franco discloses "*determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous*."

573. O'Neill discloses "*determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous*," as O'Neill discloses a system that includes what it 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:

> The DiscourseManager 22 implements the generic confirmation strategy and creates new dialogue frames for each utterance of the user by determining whether the utterance represents the repetition, modification or negation of information previously provided by the user. The DiscourseManager 22 performs this determination by comparing the dialogue frame corresponding to the user's latest utterance, with a corresponding dialogue frame representing the last discourse state taken from the DiscourseStack 123.
>
> Since the dialogue frame taken from the DiscourseStack 123 details the information previously provided by the user together with the status of the information (repeated, modified, negated, etc.) and the system's previous intentions for confirming or repairing supplied or missing information, the generic confirmation strategy enables the improved dialogue manager 20 to interpret the user's most recent utterance in light of the improved dialogue manager's own last utterance. For instance, if the user answers "Yes" in response to the system's observation "So, you're staying in Dublin", then Dublin can be regarded as the confirmed location.

(EX-1005, [0212]-[0213].)

574. Therefore, O'Neill discloses processing subsequent natural language utterances to determine if the prior natural language utterances were "*incomplete*" and then determines whether more information from the user is necessary using its confirmation strategies.

575. O'Neill further discloses the use of confirmation strategy to identify and reduce the risk of "misinterpretations of users' free-form utterances":

> Before transferring (or refusing a transfer) to a new handling agent, the improved dialogue manager always

> confirms the user's intentions ("Do you want to enquire about cinema bookings?"). Bearing in mind that the improved dialogue manager can operate in a number of different though potentially related transaction domains, the above confirmation strategy reduces the risk of serious misinterpretations of users' free-form utterances.

(EX-1005, [0236].)

576. Jones also discloses "*determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous,*" as Jones discloses seeking qualification from users when a word may have multiple meanings, i.e., is ambiguous:

> After transcribing 54 keyword(s), process 50 continues to selecting 56 advertisement data for the transcribed keyword(s). For example, advertisements associated with iPods may be selected for voice queries pertaining to apple products while advertisements associated with a particular type of the fruit apple may be selected for a query requesting information related to apple pie recipes.

(EX-1006, [0040].)

577. Jones recognizes that some words, like "apple" may have multiple meaning or connotations to a user and qualification may be necessary:

> The database 18 (FIG. 1) may maintain a histogram of usage for multiple meanings and/or forms of keyword(s) or phrase(s) which may be used to prompt a user (or consumer) or provider to qualify or clarify keyword(s) or phrase(s) having multiple meanings or forms when formulating a search query and an advertisement may be selected based on the qualified meaning of the keyword(s) or phrase(s).

(EX-1006, [0053].)

578. Jones provides an example where other promotional content may be determined and presented based on the reinterpretation or qualification:

> For example, when a search query containing the keyword "apple" is qualified to mean the fruit, an advertisement pertaining to food recipe and/or the fruit may be presented, while an advertisement pertaining to the tax program and/or tax tips are provided when the keyword "apple" in the query is qualified to mean the tax program.

(EX-1006, [0058].)

579. Franco likewise discloses "determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous."

580. Franco "relates generally to speech recognition applications and relates more particularly to correcting errors in speech recognition systems." (EX-1007, [0002].) Errors can be found when, for example, "background noise present at the time the utterance is made or speaker enunciation [] affect[s] a speech recognition application's ability to produce accurate results in a timely fashion." (EX-1007, [0004].)

581. The system in Franco assumes that "when speech recognition is unsuccessful on the first attempt, a speaker will naturally repeat all or part of the same utterance in order to correct the recognition error. Thus, acoustic evidence (e.g., recognition hypotheses) for multiple instances of the same utterance may be

combined in a manner that produces a better estimate of the utterance." (EX-1007, [0012].)

582. Franco also discloses:

> The method 200 is initialized at step 202 and proceeds to step 204, where the method 200 generates a first word lattice based on the first hypothesis. The first word lattice is a network or diagram that graphically illustrates the first hypothesis. Specifically, the first lattice may be thought of as comprising a plurality of lattice positions or word "slots", where each slot corresponds to a temporal interval in the first utterance. Each slot is associated with one or more words that may potentially be have been uttered during the corresponding temporal interval. All of the slots in total comprise a phrase that may represent the first utterance (e.g., by selecting one word for each slot and then concatenating the selected words). Each potential word is associated with a score that indicates the probability that the word was in fact, uttered during the temporal interval corresponding to that slot. In one embodiment, a deletion may be the most probable word at one or more slots in the first word lattice. In fact, if the waveform representing the first utterance has a low signal-to-noise ratio, the first word lattice may include a plurality of slots whose most probable word is a deletion, due to multiple insertions with low probabilities.

(EX-1007, [0022].)

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

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

583. The combination of O'Neill, Jones and Franco discloses "*monitoring, by the one or more physical processors, interaction of the user with the promotional content*."

584. Jones discloses monitoring user engagement with promotional content provided to the user, including items associated with the 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*":

> For example, the database 18 may maintain a record of advertisements presented in response to queries containing particular keyword(s) and whether users interacted with the advertisements.
>
> The database 18 may also maintain information associated with a number of advertisements presented while voice queries are being processed by the system 10, a number of times a particular advertisement is presented in association with one or more keywords, advertisements hand-selected by providers, etc. For example, the server 16 may compute 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*** (this information may be delivered back to the server 16 from the advertiser), etc.

(EX-1006, [0032]-[0033].)

585. O'Neill also discloses that the disclosed system 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:

> Preferably, the method includes a step of retrieving information from a previous communication stored in one or more of the data-storage units and combining the information with information from the current communication to enable the completion of a task associated with the previous communication.

(EX-1005, [0062].)

586. O'Neill further discloses:

> Furthermore, the improved dialogue manager provides a facility whereby the dialogue manager may retrieve information collected at an earlier point in a dialogue to assist in a current query.
>
> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user.

(EX-1005, [0076]-[0077].)

587. O'Neill describes the collection and use of information within a Knowledge Hierarchy:

> Referring to FIG. 2b, the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session. In particular, the knowledge-storing components store data values elicited by the system from the user, or inferred by

> the system itself, in the course of a dialogue. These components can be collectively described as the Knowledge Hierarchy.
>
> The components in the Knowledge Hierarchy also maintain information that indicates how confirmed each datum is, and what the system intends to do (if anything) to confirm that datum adequately. The data values stored in the knowledge-storing elements are the system's discourse knowledge, and are used by components of the Expertise Hierarchy as they prepare system utterances in accordance with their generic and domain specific rules.

(EX-1005, [0103]-[0104].)

588. O'Neill also discloses:

> The above-described information-storage mechanism enables the improved dialogue manager 20 to store the dialogue product (i.e. the information elicited during the evolving dialogue of a dialogue session) in a tree-like structure (henceforth known as the knowledge tree). The tree-like structure of the dialogue product emerges from the fact that a dialogue frame may include as part of its Attributes, links to dialogue frames of other types (e.g. TheatreDialogFrame 133 may include a link to a PaymentDialogFrame 152).

(EX-1005, [0184].)

589. Therefore, a POSA would be motivated to combinate the user tracking disclosed by O'Neill with Jones and create a system that also tracks the user's interactions with promotional content, including items associated with the promotional content. The system disclosed by O'Neill tracks all user interactions with the system so it would be automatic for the O'Neill system to similarly track user interactions with promotional content.

590. As discussed above, O'Neill and Jones both perform the disclosed methods using "one or more physical processors," including "*monitoring, by the one or more physical processors, interaction of the user with the promotional content*." (§V.B.1.)

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

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

591. The combination of O'Neill, Jones and Franco discloses "*interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction*."

592. As discussed above, O'Neill discloses that the system monitors the user's interactions with the system and the system's interactions with the user to improve the dialogue with the user, which in the system combined with Jones would include user "*interaction*." O'Neill discloses that the dialogue manager continues to use information from prior parts of the conversation with the user, which would include user "*interaction*":

> Preferably, the method includes a step of retrieving information from a previous communication stored in one or more of the data-storage units and combining the information with information from the current communication to enable the completion of a task associated with the previous communication.

(EX-1005, [0062].)

593. O'Neill further discloses:

> Furthermore, the improved dialogue manager provides a facility whereby the dialogue manager may retrieve information collected at an earlier point in a dialogue to assist in a current query.
>
> The above features of the improved dialogue manager enable a discourse structure and corresponding dialogue product to evolve dynamically, as agents are selected in light of the user's utterances or as a consequence of the agents' own rules. It is this process, rather than an over-arching dialogue plan or agenda that drives the discourse forward, and across domain boundaries if required by the user.

(EX-1005, [0076]-[0077].)

594. O'Neill describes the collection and use of information within a Knowledge Hierarchy:

> Referring to FIG. 2 b, the improved dialogue manager further comprises components for storing information acquired or generated during a dialogue session. In particular, the knowledge-storing components store data values elicited by the system from the user, or inferred by the system itself, in the course of a dialogue. These components can be collectively described as the Knowledge Hierarchy.
>
> The components in the Knowledge Hierarchy also maintain information that indicates how confirmed each datum is, and what the system intends to do (if anything) to confirm that datum adequately. The data values stored in the knowledge-storing elements are the system's discourse knowledge, and are used by components of the Expertise Hierarchy as they prepare system utterances in accordance with their generic and domain specific rules.

(EX-1005, [0103]-[0104].)

595. O'Neill also discloses:

> The above-described information-storage mechanism enables the improved dialogue manager 20 to store the dialogue product (i.e. the information elicited during the evolving dialogue of a dialogue session) in a tree-like structure (henceforth known as the knowledge tree). The tree-like structure of the dialogue product emerges from the fact that a dialogue frame may include as part of its Attributes, links to dialogue frames of other types (e.g. TheatreDialogFrame 133 may include a link to a PaymentDialogFrame 152).

(EX-1005, [0184].)

596. Therefore, a POSA would be motivated to combine the tracking of user "*interaction*" disclosed by Jones with the dialogue system disclosed in O'Neill that would use user "*interaction*" to interpret "*incomplete or ambiguous*" utterances. The system disclosed by O'Neill tracks all user interactions with the system, including "*interactions*" so it would be automatic for the O'Neill system to similarly use interactions in its dialogue processor.

## F. Independent Claim 33

### 1. Preamble [33P]

| |
|---|
| [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: |

597. To the extent the preamble is limiting, the combination of O'Neill, Jones and Franco discloses "*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*."

598. The preamble of claim 33 is substantially the same limitation that is addressed in §V.B.1 regarding claim limitation [1P] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

599. For the same reasons discussed above in §V.B.1 for claim limitation [1P], the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Receiving a First Natural Language Utterance [33A]

| |
|---|
| [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; |

600. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a first natural language utterance during an interactive session between a user and the computer system*."

601. This limitation of claim 33 is substantially the same limitation that is addressed in §V.C.1 regarding claim limitation [2A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

602. For the same reasons discussed above in §V.C.1 for claim limitation [2A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 3. Providing a Response to the First Natural Language Utterance [33B]

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

603. The combination of O'Neill, Jones and Franco discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

604. This limitation of claim 33 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [33B] providing, by the one or more physical processors, a response to the first natural language utterance; |

605. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill, Jones and Franco discloses this limitation.

#### 4. Receiving a Second Natural Language Utterance [33C]

| |
|---|
| [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; |

606. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance during the interactive session*."

607. This limitation of claim 33 is substantially the same limitation that is addressed in §§V.B.4; V.C.1 regarding claim limitations [1C] and [2A] above. I provide a side-by-side comparison of the limitations here:

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

608. For the same reasons discussed above in §§V.B.4; V.C.1 for claim limitations [1C] and [2A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 5. Performing Speech Recognition to Recognize Words of Second Natural Language Utterance [33D]

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

609. The combination of O'Neill, Jones and Franco discloses "*performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance*."

610. This limitation of claim 33 is substantially the same limitation that is addressed in §V.B.5 regarding claim limitation [1D] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance; | [33D] performing, by the one or more physical processors, speech recognition to recognize one or more words of the second natural language utterance; |

611. For the same reasons discussed above in §V.B.5 for claim limitation [1D], the combination O'Neill, Jones and Franco discloses this limitation.

### 6. Determining an Interpretation of the Second Natural Language Utterance [33E]

| |
|---|
| [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; |

612. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance during the interactive session*."

613. This limitation of claim 33 is substantially the same limitation that is addressed in §V.B.7 regarding claim limitation [1F] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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; | [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; |

614. For the same reasons discussed above in §V.B.7 for claim limitation [1F], the combination of O'Neill, Jones and Franco discloses this limitation.

615. The only difference between claim limitations [1F] and [33E] is that claim limitation [1F] processes "*based on the domain information*" while claim limitation [33E] processes "*based on the first natural language utterance*." However, as discussed above, the domain information used to process the second

natural language utterance in claim 1 is based on the first natural language utterance. (*See supra* §V.B.7.)

### 7. Receiving a Third Natural Language Utterance [33F]

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

616. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a third natural language utterance during the interactive session*."

617. This limitation of claim 33 is substantially the same limitation that is addressed in §VI.B.2 regarding claim limitation [13B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [13B] receiving, at the one or more physical processors, a third natural language utterance during the interactive session; | [33F] receiving, at the one or more physical processors, a third natural language utterance during the interactive session; |

618. For the same reasons discussed above in §VI.B.2 for claim limitation [13B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 8. Processing the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [33G]

| |
|---|
| [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; |

619. The combination of O'Neill, Jones and Franco discloses "*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*."

620. This limitation of claim 33 is substantially the same limitation that is addressed in §VI.B.3 regarding claim limitation [13C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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. | [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; |

621. For the same reasons discussed above in §VI.B.3 for claim limitation [13C], the combination of O'Neill, Jones and Franco discloses this limitation.

### 9. Determining Promotional Content [33H]

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

622. The combination of O'Neill, Jones and Franco discloses "*determining, by the one or more physical processors, promotional content based on one or more of the interpretation or the reinterpretation*."

623. This limitation of claim 33 is substantially the same limitation that is addressed in §V.B.11 regarding claim limitations [1K] (interpretation) and [14] (reinterpretation) above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1K] determining, by the one or more physical processors, promotional content based on the interpretation; and<br><br>[14] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the reinterpretation; | [33H] determining, by the one or more physical processors, promotional content based on one or more of the interpretation or the reinterpretation; and |

624. For the same reasons discussed above in §V.B.11 for claim limitations [1K] and [14], the combination of O'Neill, Jones and Franco discloses this limitation.

625. The difference between claim 14 and claim limitation [33H] is that claim 14 determined promotion content "based on the reinterpretation" of the second natural language utterance while claim limitation [33H] determines promotion content "*based on one or more of the interpretation* ***or*** *the reinterpretation*." As claim limitation [33H] requires the determination to be based on "*the interpretation*" **or** "*the reinterpretation*" the analysis above in §V.B.11 for claim limitation 14 is applicable to "*the reinterpretation*."

### 10. Receiving a Second Natural Language Utterance [33I]

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

626. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance during the interactive session*."

627. This limitation of claim 33 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [33I] presenting, by the one or more physical processors, the promotional content to the user. |

628. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill, Jones and Franco discloses this limitation.

### G. Dependent Claim 34

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

629. The combination of O'Neill, Jones and Franco discloses "*wherein determining the promotional content comprises determining the promotional content based on the reinterpretation*."

630. This limitation of claim 34 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [14] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [14] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the reinterpretation. | [34] The method of claim 33, wherein determining the promotional content comprises determining the promotional content based on the reinterpretation. |

631. For the same reasons discussed above in §VI.C for claim limitation [14], the combination of O'Neill, Jones and Franco discloses this limitation.

## H. Dependent Claim 35

### 1. Determining the Promotional Content Based on the Interpretation [35A]

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

632. The combination of O'Neill, Jones and Franco discloses "*wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising*."

633. This limitation of claim 35 is substantially the same limitation that is addressed in §VI.D.1 regarding claim limitation [15A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [15A] The method of claim 13, wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising: | [35A] The method of claim 33, wherein determining the promotional content comprises determining the promotional content based on the interpretation, the method further comprising: |

634. For the same reasons discussed above in §VI.D.1 for claim limitation [15A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Determining Other Promotional Content Based on the Interpretation [35B]

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

635. The combination of O'Neill, Jones and Franco discloses "*determining, by the one or more physical processors, other promotional content based on the reinterpretation*."

636. This limitation of claim 35 is substantially the same limitation that is addressed in §VI.D.2 regarding claim limitation [15B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [15B] determining, by the one or more physical processors, other promotional content based on the reinterpretation; and | [35B] determining, by the one or more physical processors, other promotional content based on the reinterpretation; and |

637. For the same reasons discussed above in §VI.D.2 for claim limitation [15B], the combination of O'Neill, Jones and Franco discloses this limitation.

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

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

638. The combination of O'Neill, Jones and Franco discloses "*presenting, by the one or more physical processors, the other promotional content to the user*."

639. This limitation of claim 35 is substantially the same limitation that is addressed in §VI.D.3 regarding claim limitation [15C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [15C] presenting, by the one or more physical processors, the other promotional content to the user. | [35C] presenting, by the one or more physical processors, the other promotional content to the user. |

640. For the same reasons discussed above in §VI.D.3 for claim limitation [15C], the combination O'Neill, Jones and Franco discloses this limitation.

## I. Independent Claim 36

### 1. Preamble [36P]

| |
|---|
| [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: |

641. To the extent this preamble is limiting, the combination of O'Neill, Jones and Franco discloses "*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*."

642. This preamble of claim 36 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 32 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

643. For the same reasons discussed above in §V.B.1 for the preamble of claim 32, the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Receive a First Natural Language Utterance [36A]

| |
|---|
| [36A] receive a first natural language utterance; |

644. The combination of O'Neill, Jones and Franco discloses "*receive a first natural language utterance*."

645. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [32A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32A] receive a first natural language utterance; | [36A] receive a first natural language utterance; |

646. For the same reasons discussed above in §V.B.2 for claim limitation [32A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 3. Provide Response to First Natural Language Utterance [36B]

| |
|---|
| [36B] provide a response to the first natural language utterance; |

647. The combination of O'Neill, Jones and Franco discloses "*provide a response to the first natural language utterance*."

648. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [32B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32B] provide a response to the first natural language utterance; | [36B] provide a response to the first natural language utterance; |

649. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill, Jones and Franco discloses this limitation.

#### 4. Receive a Second Natural Language Utterance [36C]

| |
|---|
| [36C] receive a second natural language utterance relating to the first natural language utterance; |

650. The combination of O'Neill, Jones and Franco discloses "*receive a second natural language utterance relating to the first natural language utterance*."

651. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [32C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32C] receive a second natural language utterance relating to the first natural language utterance; | [43C] receive a second natural language utterance relating to the first natural language utterance; |

652. For the same reasons discussed above in §V.B.4 for claim limitation [32C], the combination of O'Neill, Jones and Franco discloses this limitation.

### 5. Perform Speech Recognition to Recognize One or More Words in Second Natural Utterance [36D]

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

653. The combination of O'Neill, Jones and Franco discloses "*perform speech recognition to recognize one or more words of the second natural language utterance*."

654. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.5 regarding claim limitation [32D] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32D] perform speech recognition to recognize one or more words of the second natural language utterance; | [36D] perform speech recognition to recognize one or more words of the second natural language utterance; |

655. For the same reasons discussed above in §V.B.5 for claim limitation [32D], the combination of O'Neill, Jones and Franco discloses this limitation.

### 6. Process One or More Recognized Words to Determine Interpretation of Second Natural Language Utterance [36E]

| |
|---|
| [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; |

656. The combination of O'Neill, Jones and Franco discloses "*process, based on the first natural language utterance, the one or more recognized words to determine an interpretation of the second natural language utterance*."

657. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.7 regarding claim limitation [1F] above. I provide a side-by-side comparison of the limitations here:

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

658. For the same reasons discussed above in §V.B.7 for claim limitation [1F], the combination of O'Neill, Jones and Franco discloses this limitation. The only difference between claim limitation [1F] and [36D] is that claim limitation [1F] requires "*processing ... based on the domain information*" while claim limitation [36D] requires the system to "*process ... based on the first natural language utterance.*" However, as discussed above in claim 1 §V.B.7, the "*domain information*" is determined based on the first natural language utterance.

### 7. Receive a Third Natural Language Utterance [36F]

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

659. The combination of O'Neill, Jones and Franco discloses "*receive a third natural language utterance during the interactive session.*"

660. This limitation of claim 36 is substantially the same limitation that is addressed in §VI.B.2 regarding claim limitation [13B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [13B] receiving, at the one or more physical processors, a third natural language utterance during the interactive session; | [36F] receive a third natural language utterance during the interactive session; |

661. For the same reasons discussed above in §VI.B.2 for claim limitation [13B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 8. Process the Third Natural Language Utterance to Determine a Reinterpretation of the Second Natural Language Utterance [36G]

| |
|---|
| [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; |

662. The combination of O'Neill, Jones and Franco discloses "*process, based on the third natural language utterance, the one or more recognized words to determine a reinterpretation of the second natural language utterance*."

663. This limitation of claim 36 is substantially the same limitation that is addressed in §VI.B.3 regarding claim limitation [13C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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. | [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; |

664. For the same reasons discussed above in §VI.B.3 for claim limitation [13C], the combination of O'Neill, Jones and Franco discloses this limitation.

### 9. Determine Promotional Content Based on One or More of the Interpretation or the Reinterpretation [36H]

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

665. The combination of O'Neill, Jones and Franco discloses "*determine promotional content based on one or more of the interpretation or the reinterpretation*."

666. This limitation of claim 36 is substantially the same limitation that is addressed in §VI.F.9 regarding claim limitation [33H] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [33H] determining, by the one or more physical processors, promotional content based on one or more of the interpretation or the reinterpretation; and | [36H] determine promotional content based on one or more of the interpretation or the reinterpretation; and |

667. For the same reasons discussed above in §VI.F.9 for claim limitation [33H], the combination of O'Neill, Jones and Franco discloses this limitation.

### 10. Present the Promotional Content to the User [36I]

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

668. The combination of O'Neill, Jones and Franco discloses "*present the promotional content to a user*."

669. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [32L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32L] present the promotional content to a user. | [36I] present the promotional content to a user. |

670. For the same reasons discussed above in §V.B.12 for claim limitation [32L], the combination of O'Neill, Jones and Franco discloses this limitation.

## J. Independent Claim 37

### 1. Preamble [37P]

| |
|---|
| [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: |

671. To the extent the preamble is limiting, the combination of O'Neill, Jones and Franco discloses "*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*."

672. The preamble of claim 37 is substantially the same limitation that is addressed in §V.B.1 regarding claim limitation [1P] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

673. For the same reasons discussed above in §V.B.1 for claim limitation [1P], the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Receiving a First Natural Language Utterance [37A]

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

674. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a first natural language utterance*."

675. This limitation of claim 37 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [37A] receiving, at the one or more physical processors, a first natural language utterance; |

676. For the same reasons discussed above in § V.B.2 for claim limitation [1A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 3. Providing a Response to the First Natural Language Utterance [37B]

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

677. The combination of O'Neill, Jones and Franco discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

678. This limitation of claim 37 is substantially the same limitation that is addressed in § V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [37B] providing, by the one or more physical processors, a response to the first natural language utterance; |

679. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 4. Receiving a Second Natural Language Utterance [37C]

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

680. The combination of O'Neill, Jones and Franco discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*."

681. This limitation of claim 37 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

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

682. For the same reasons discussed above in §V.B.4 for claim limitation [1C], the combination of O'Neill, Jones and Franco discloses this limitation.

### 5. Identifying a Request [37D]

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

683. The combination of O'Neill, Jones and Franco discloses "identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance."

684. This limitation of claim 37 is substantially the same limitation that is addressed in §V.N.2 regarding claim limitation [16B] above. I provide a side-by-side comparison of the limitations here:

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

685. For the same reasons discussed above in §V.N.2 for claim limitation [16B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 6. Incomplete or Ambiguous Request [37E]

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

686. The combination of O'Neill, Jones and Franco discloses "determining, by the one or more physical processors, that at least one request of the one or more requests is incomplete or ambiguous."

687. This limitation of claim 37 is substantially the same limitation that is addressed in §VI.E.1 regarding claim limitation [24A] above. I provide a side-by-side comparison of the limitations here:

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

688. For the same reasons discussed above in §VI.E.1 for claim limitation [24A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 7. Determining Promotional Content Related to Request [37F]

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

689. The combination of O'Neill, Jones and Franco discloses "determining, by the one or more physical processors, promotional content that relates to the one or more requests."

690. This limitation of claim 37 is substantially the same limitation that is addressed in §V.B.11 regarding claim limitation [1K] above. 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. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1K] determining, by the one or more physical processors, promotional content based on the interpretation; and | [37F] determining, by the one or more physical processors, promotional content that relates to the one or more requests; |

691. For the same reasons discussed above in §V.B.11 for claim limitation [1K], the combination of O'Neill, Jones and Franco discloses this limitation.

### 8. Present Promotional Content [37G]

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

692. The combination of O'Neill, Jones and Franco discloses "presenting, by the one or more physical processors, the promotional content to a user."

693. This limitation of claim 37 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [37G] presenting, by the one or more physical processors, the promotional content to a user |

694. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill, Jones and Franco discloses this limitation.

### 9. Identifying a Request [37H]

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

695. The combination of O'Neill, Jones and Franco discloses "monitoring, by the one or more physical processors, interaction of the user with the promotional content."

696. This limitation of claim 37 is substantially the same limitation that is addressed in §VI.E.2 regarding claim limitation [24B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [24B] monitoring, by the one or more physical processors, interaction of the user with the promotional content; and | [37H] monitoring, by the one or more physical processors, interaction of the user with the promotional content; and |

697. For the same reasons discussed above in §VI.E.2 for claim limitation [24B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 10. Incomplete or Ambiguous Request Based on the Interaction [37I]

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

698. The combination of O'Neill, Jones and Franco discloses "interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction."

699. This limitation of claim 37 is substantially the same limitation that is addressed in §VI.E.3 regarding claim limitation [24C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [24C] interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction. | [37I] interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction. |

700. For the same reasons discussed above in §VI.E.3 for claim limitation [24C], the combination of O'Neill, Jones and Franco discloses this limitation.

## K. Independent Claim 38

### 1. Preamble [38P]

| |
|---|
| [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: |

701. To the extent this preamble is limiting, the combination of O'Neill, Jones and Franco discloses "*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*."

702. This preamble of claim 38 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 32 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

703. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill, Jones and Franco discloses this limitation.

### 2. Receive a First Natural Language Utterance [38A]

| |
|---|
| [38A] receive a first natural language utterance; |

704. The combination of O'Neill, Jones and Franco discloses "*receive a first natural language utterance*."

705. This limitation of claim 36 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [32A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32A] receive a first natural language utterance; | [38A] receive a first natural language utterance; |

706. For the same reasons discussed above in §V.B.2 for claim limitation [32A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 3. Provide Response to First Natural Language Utterance [38B]

| |
|---|
| [38B] provide a response to the first natural language utterance; |

707. The combination of O'Neill, Jones and Franco discloses "*provide a response to the first natural language utterance*."

708. This limitation of claim 38 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [32B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32B] provide a response to the first natural language utterance; | [38B] provide a response to the first natural language utterance; |

709. For the same reasons discussed above in §V.B.3 for claim limitation [32B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 4. Receive a Second Natural Language Utterance [36C]

| |
|---|
| [38C] receive a second natural language utterance relating to the first natural language utterance; |

710. The combination of O'Neill, Jones and Franco discloses "*receive a second natural language utterance relating to the first natural language utterance*."

711. This limitation of claim 38 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [32C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32C] receive a second natural language utterance relating to the first natural language utterance; | [38C] receive a second natural language utterance relating to the first natural language utterance; |

712. For the same reasons discussed above in §V.B.4 for claim limitation [32C], the combination of O'Neill, Jones and Franco discloses this limitation.

### 5. Identify One or More Requests Associated with the Second Natural Language Utterance [38D]

| |
|---|
| [38D] identify one or more requests associated with the second natural language utterance; |

713. The combination of O'Neill, Jones and Franco discloses "*identify one or more requests associated with the second natural language utterance*."

714. This limitation of claim 38 is substantially the same limitation that is addressed in §V.B.5 regarding claim limitation [16A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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, | [38D] identify one or more requests associated with the second natural language utterance; |

715. For the same reasons discussed above in §V.B.5 for claim limitation [16A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 6. Determine at Least One Request is Incomplete or Ambiguous [38E]

| |
|---|
| [38E] determine that at least one request of the one or more requests is incomplete or ambiguous; |

716. The combination of O'Neill, Jones and Franco discloses "*identify one or more requests associated with the second natural language utterance*."

717. This limitation of claim 38 is substantially the same limitation that is addressed in §VI.E.1 regarding claim limitation [24A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [38E] determine that at least one request of the one or more requests is incomplete or ambiguous; |

718. For the same reasons discussed above in §VI.E.1 for claim limitation [24A], the combination of O'Neill, Jones and Franco discloses this limitation.

### 7. Determine Promotional Content Relating to Request [38F]

| |
|---|
| [38F] determine promotional content that relates to the one or more requests; and |

719. The combination of O'Neill, Jones and Franco discloses "*determine promotional content that relates to one or more of the first request or the second request*."

720. This limitation of claim 38 is substantially the same limitation that is addressed in §§V.B.11; VI.J.7 regarding claim limitations [1K] and [37F] above. I provide a side-by-side comparison of limitations [37F] and [38F] here:

| | |
|---|---|
| [37F] determining, by the one or more physical processors, promotional content that relates to the one or more requests; | [38F] determine promotional content that relates to the one or more requests; and |

721. For the same reasons discussed above in §§V.B.11; VI.J.7 for claim limitations [1K]/[37F], the combination of O'Neill, Jones and Franco discloses this limitation.

### 8. Present the Promotional Content to the User [38G]

| |
|---|
| [38G] present the promotional content to the user. |

722. The combination of O'Neill, Jones and Franco discloses "*present the promotional content to a user*."

723. This limitation of claim 38 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [32L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [32L] present the promotional content to a user. | [38G] present the promotional content to a user. |

724. For the same reasons discussed above in §V.B.12 for claim limitation [32L], the combination of O'Neill, Jones and Franco discloses this limitation.

### 9. Monitor Interactions of the User with the Promotional Content [38H]

| |
|---|
| [38H] monitor interaction of the user with the promotional content; and |

725. The combination of O'Neill, Jones and Franco discloses "*present the promotional content to a user*."

726. This limitation of claim 38 is substantially the same limitation that is addressed in §VI.E.2 regarding claim limitation [24B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [24B] monitoring, by the one or more physical processors, interaction of the user with the promotional content; and | [38H] monitor interaction of the user with the promotional content; and |

727. For the same reasons discussed above in §VI.E.2 for claim limitation [24B], the combination of O'Neill, Jones and Franco discloses this limitation.

### 10. Interpret the Incomplete or Ambiguous Request Based on the Interaction [38I]

| |
|---|
| [38I] interpret the at least one incomplete or ambiguous request based on the interaction. |

728. The combination of O'Neill, Jones and Franco discloses "*present the promotional content to a user*."

729. This limitation of claim 38 is substantially the same limitation that is addressed in §VI.E.3 regarding claim limitation [24C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [24C] interpreting, by the one or more physical processors, the at least one incomplete or ambiguous request based on the interaction. | [38I] interpret the at least one incomplete or ambiguous request based on the interaction |

730. For the same reasons discussed above in §VI.E.3 for claim limitation [24C], the combination of O'Neill, Jones and Franco discloses this limitation.

## I. GROUND 3: THE COMBINATION OF O'NEILL, JONES AND WANG RENDERS OBVIOUS CLAIMS 25-26, 31 AND 39-43

731. The combination of O'Neill, Jones and U.S. Patent No. 7,609,829 ("Wang"; EX-1008) discloses every limitation of claims 25-26, 31 and 39-43. O'Neill, Jones and Wang are analogous art to the '536 patent because all three references are directed to speech recognition systems and software.

732. I have described both O'Neill and Jones above, and I discuss Wang below.

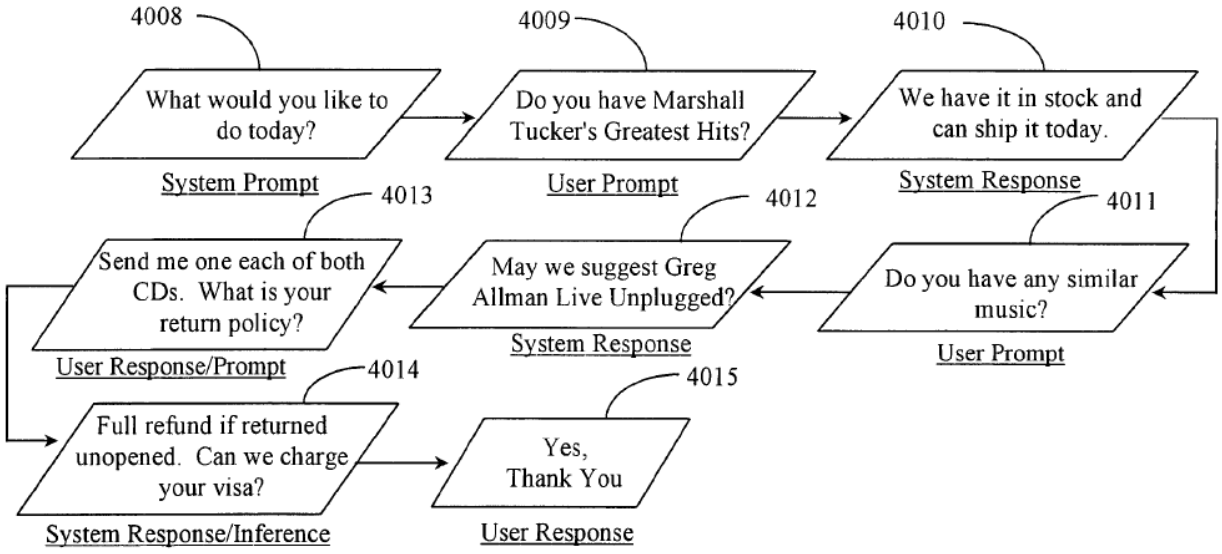
### A. Combination Overview

#### 1. Wang

733. 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.)

734. "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.)

735. 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

736. A POSITA would have been motivated to combine the functionality for processing multi-request utterances disclosed in Wang into the O'Neill-Jones combination for a number of reasons.

737. 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.

738. 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).

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

### B. Dependent Claim 25

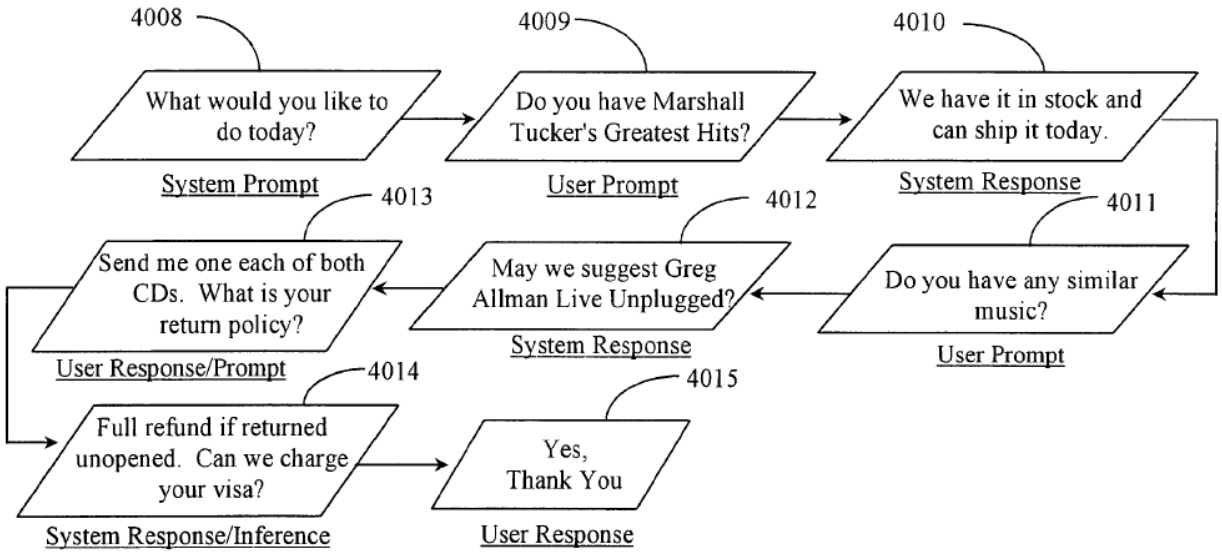
#### 1. First Request Associated with First Application and Second Request Associated with Second Application [25A]

| |
|---|
| [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 |

740. The combination of O'Neill, Jones and Wang discloses "*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*."

741. Wang discloses a single utterance containing a "*first request*" and a "*second request*." 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**

742. 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.

743. Wang further explains that "[a] user may call or navigate to server 3408 for interaction." (EX-1008, 50:50.) Wang also discloses that "[a]pplication servers

3406 and 3428 represent any hosted servers adapted to allow 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.)

744. Following the same example, Wang explains that the "order preparation can take place in the background while the system addresses the second part of the response 'What is your 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. 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, for example: "Show sales in a list, and represent its contents by a bar graph." (EX-1013, 23:17-18.) Namba explains that the 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.)

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

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

745. The combination of O'Neill, Jones and Wang discloses "*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 wherein the promotional content relates to one or more of the request or the second request*."

746. This limitation of claim 25 is substantially the same limitation that is addressed in §V.B.11 regarding claim limitation [1K] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1K] determining, by the one or more physical processors, promotional content based on the interpretation; and | [25B] wherein the promotional content relates to one or more of the first request or the second request. |

747. For the same reasons discussed above in §V.B.11 for claim limitation [1K], the combination of O'Neill, Jones and Wang discloses this limitation.

### C. Dependent Claims 26 and 40

[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.

[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.

748. The combination of O'Neill, Jones and Wang discloses "*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*."

749. Wang discloses a multi-request utterance during 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.)

## D. Dependent Claim 31

### 1. Obtaining User Profile Information [31A]

| |
|---|
| [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; |

750. The combination of O'Neill, Jones and Wang discloses "*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*."

751. This limitation of claim 31 is substantially the same limitation that is addressed in §V.V.1 regarding claim limitation [30A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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; |

752. For the same reasons discussed above in §V.V.1 for claim limitation [30A], the combination of O'Neill, Jones and Wang discloses this limitation.

### 2. Identifying One or More Requests Associated with the First or Second Natural Language Utterance [31B]

| |
|---|
| [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; |

753. The combination of O'Neill, Jones and Wang discloses "*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*."

754. This limitation of claim 31 is substantially the same limitation that is addressed in §V.N.2 regarding claim limitation [16B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [16B] identifying, by the one or more physical processors, one or more requests associated with the second natural language utterance, | [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; |

755. While claim limitation [16B] requires identifying "*one or more requests associated with the second natural language utterance*," claim limitation [31B] requires identifying "*one or more requests associated with the first natural language utterance* ***or*** *the second natural language utterance*." Therefore, for the same reasons discussed above in §V.N.2 for claim limitation [16B], the combination of O'Neill, Jones and Wang discloses this limitation.

756. Moreover, O'Neill also discloses identifying requests associated with the first natural language utterance. For example, O'Neill discloses an example where a user's first natural language utterance is "I'd like to go see THE LION KING 2," and in response the system identifies that the user is making a request for

assistance booking a cinema ticket and use the CinemaExpert domain agent to process the user's request. (EX-1005, [0203]-[0204].)

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

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

757. The combination of O'Neill, Jones and Wang discloses "*determining, by the one or more physical processors, one or more applications for processing the one or more requests*."

758. As discussed above in §VII.B.1, O'Neill and Wang associate user requests with applications. Limitation [31B] additionally requires determining the application for "*processing*" the request. O'Neill and Wang disclose this limitation because 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 §VII.B.1, the combination of O'Neill, Jones and Wang combination discloses these limitations.

759. O'Neill provides an example, discussed above, where a user first asks the system to assist in the booking of a hotel accommodation. The O'Neill system processes that request using the domain specific AccomodationExpert, an "*application*" which assists the user with the hotel booking. (EX-1005, [0123].)

760. During the same interactive dialogue session with the user, the user also requests assistance booking a cinema ticket. The O'Neill system transfers the user to the CinemaExpert, another "*application*," to assist the user with the booking of the cinema ticket. (EX-1005, [0290]-[0408].)

761. O'Neill further discloses that the demands of specific applications will determine how the system processes the user requests:

> In the present example, the class of service agents 30 include an AccommodationExpert 31 and an EventExpert 32 that are respectively responsible for searching for and booking accommodation and event tickets. In addition, the class of support agents 50 include an AddressExpert 51 and a PaymentExpert 52 that are respectively responsible for acquiring the user's address and payment details. It 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, [0123].)

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

| |
|---|
| [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, |

762. The combination of O'Neill, Jones and Wang discloses "*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*."

763. This limitation of claim 31 is substantially the same limitation that is addressed in §V.V.2 regarding claim limitation [30B] above. I provide a side-by-side comparison of the limitations here:

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

764. For the same reasons discussed above in §V.V.2 for claim limitation [30B], the combination of O'Neill, Jones and Wang discloses this limitation.

765. The only portion of this claim limitation that does not appear in claim limitation [30B] is the portion "*wherein the categories relate to the one or more applications*." That portion of this claim limitation is also disclosed in Jones. For example, Jones discloses that:

> Using the same example in the previous discussion, if a user interacts with the advertisement pertaining to the specific discounted Dell products, an option to receive a coupon for the products may be provided to the user (see also FIG. 3B). In the case of electronic merchandise (eg. MP3 downloads) a product may be delivered in addition to an advertisement. For example the advertisement may contain video or audio 'clips' ad the consumer may elect to purchase the full-length product, which is then supplied to him.

(EX-1006, [0055].) In this example, the system identified 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.

766. Similarly, O'Neill discloses "*categories of items based on the prior user interactions*" as the system assist the user book 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].) In O'Neill the AcommodationExpert handles the hotel accommodation request and the CinemaExpert handles the cinema ticket request. (EX-1005, [0294], [0394].) While O'Neill does not disclose using this user profile information to determine promotional content, when O'Neill is combined with Jones the combination would determine promotional content and promotional items based on the categories disclosed in O'Neill.

### 5. Determining Promotional Item Associated with One or More Categories [31E]

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

767. The combination of O'Neill, Jones and Wang discloses "*wherein determining the promotional content comprises determining a promotional item associated with one of the categories*."

768. This limitation of claim 31 is substantially the same limitation that is addressed in §V.V.3 regarding claim limitation [30C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [30C] wherein determining the promotional content comprises determining a promotional item associated with one of the categories. | [31E] wherein determining the promotional content comprises determining a promotional item associated with one of the categories. |

769. For the same reasons discussed above in §V.V.C for claim limitation [30C], the combination of O'Neill, Jones and Wang discloses this limitation.

## E. Independent Claim 39

### 1. Preamble [39P]

| |
|---|
| [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: |

770. To the extent the preamble is limiting, the combination of O'Neill, Jones and Wang discloses "*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*."

771. The preamble of claim 39 is substantially the same limitation that is addressed in §V.B.1 regarding claim limitation [1P] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

772. For the same reasons discussed above in §V.B.1 for claim limitation [1P], the combination of O'Neill, Jones and Wang discloses this limitation.

### 2. Receiving a First Natural Language Utterance [39A]

| |
|---|
| [39A] receiving, at the one or more physical processors, a first natural language utterance; |

773. The combination of O'Neill, Jones and Wang discloses "*receiving, at the one or more physical processors, a first natural language utterance*."

774. This limitation of claim 39 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [39A] receiving, at the one or more physical processors, a first natural language utterance; |

775. For the same reasons discussed above in §V.B.2 for claim limitation [1A], the combination of O'Neill, Jones and Wang discloses this limitation.

### 3. Providing a Response to the First Natural Language Utterance [39B]

| |
|---|
| [39B] providing, by the one or more physical processors, a response to the first natural language utterance; |

776. The combination of O'Neill, Jones and Wang discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

777. This limitation of claim 39 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [39B] providing, by the one or more physical processors, a response to the first natural language utterance; |

778. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill, Jones and Wang discloses this limitation.

### 4. Receiving a Second Natural Language Utterance [39C]

| |
|---|
| [39C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance; |

779. The combination of O'Neill, Jones and Wang discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*."

780. This limitation of claim 39 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

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

781. For the same reasons discussed above in §V.B.4 for claim limitation [1C], the combination of O'Neill, Jones and Wang discloses this limitation.

### 5. Identifying Requests Associated With Applications [39D] / Determining Promotional Content Relating to Request [39E]

| |
|---|
| [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;<br><br>[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 |

782. The combination of O'Neill, Jones and Wang discloses "*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" and "determining, by the one or more physical*

*processors, promotional content that relates to one or more of the first request or the second request*."

783. These limitations of claim 39 is substantially the same limitation that is addressed in §VII.B regarding claim limitations [25A]-[25B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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<br><br>[25B] wherein the promotional content relates to one or more of the first request or the second request. | [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;<br><br>[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 |

784. For the same reasons discussed above in §VII.B. for claim limitations [25A]-[25B], the combination of O'Neill, Jones and Wang discloses these limitations.

### 6. Presenting Promotional Content [39F]

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

785. The combination of O'Neill, Jones and Wang discloses "*presenting, by the one or more physical processors, the promotional content to a user*."

786. This limitation of claim 39 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [39F] presenting, by the one or more physical processors, the promotional content to a user. |

787. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill, Jones and Wang discloses this limitation.

### F. Dependent Claim 41

> [41] The method of claim 39, further comprising: receiving, at the one or more physical processors, a third natural language utterance or the second natural language utterance; and determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content.

788. The combination of O'Neill, Jones and Wang discloses "*receiving, at the one or more physical processors, a third natural language utterance or the second natural language utterance; and determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content*."

789. This limitation of claim 41 is substantially the same limitation that is addressed in §V.H regarding claim limitation [7A]-[7B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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<br><br>[7B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content. | [41] The method of claim 39, further comprising: receiving, at the one or more physical processors, a third natural language utterance or the second natural language utterance; and determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content. |

790. For the same reasons discussed above in §V.H for claim limitation [7A]-[7B], the combination of O'Neill, Jones and Wang discloses this limitation.

### G. Dependent Claim 42

| |
|---|
| [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. |

791. The combination of O'Neill, Jones and Wang discloses "*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*."

792. This limitation of claim 41 is substantially the same limitation that is addressed in §V.M regarding claim limitation [12] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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. | [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. |

793. For the same reasons discussed above in §V.M for claim limitation 12, the combination of O'Neill, Jones and Wang discloses this limitation.

## H. Independent Claim 43

### 1. Preamble [43P]

| |
|---|
| [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: |

794. To the extent this preamble is limiting, the combination of O'Neill, Jones and Wang discloses *"[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*."

795. This preamble of claim 43 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claim 1 above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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: | [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: |

796. For the same reasons discussed above in §V.B.1 for the preamble of claim 1, the combination of O'Neill, Jones and Wang discloses this limitation.

### 2. Provide Response to First Natural Language Utterance [43B]

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

797. The combination of O'Neill, Jones and Wang discloses "*provide a response to the first natural language utterance*."

798. This limitation of claim 43 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [43B] provide a response to the first natural language utterance; |

799. For the same reasons discussed above in §V.B.3 for claim limitation [1B], the combination of O'Neill, Jones and Wang discloses this limitation.

### 3. Receive a Second Natural Language Utterance [43C]

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

800. The combination of O'Neill, Jones and Wang discloses "*receive a second natural language utterance relating to the first natural language utterance*."

801. This limitation of claim 43 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitation [1C] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1C] receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance; | [43C] receive a second natural language utterance relating to the first natural language utterance; |

802. For the same reasons discussed above in §V.B.4 for claim limitation [1C], the combination of O'Neill, Jones and Wang discloses this limitation.

#### 4. Identify Requests Associated With Applications [43D]

| |
|---|
| [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; |

803. The combination of O'Neill, Jones and Wang discloses "*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*."

804. This limitation of claim 43 is substantially the same limitation that is addressed in §V.B.5 regarding claim limitation [39D] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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; | [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; |

805. For the same reasons discussed above in §V.B.5 for claim limitation [39D], the combination of O'Neill, Jones and Wang discloses this limitation.

### 5. Determine Promotional Content Relating to Request [43E]

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

806. The combination of O'Neill, Jones and Wang discloses "*determine promotional content that relates to one or more of the first request or the second request*."

807. This limitation of claim 43 is substantially the same limitation that is addressed in §VII.E.5 regarding claim limitation [39E] above. I provide a side-by-side comparison of the limitations here:

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

808. For the same reasons discussed above in §VII.E.5 for claim limitation [39E], the combination of O'Neill, Jones and Wang discloses this limitation.

### 6. Present Promotional Content [43F]

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

809. The combination of O'Neill, Jones and Wang discloses "*present the promotional content to a user*."

810. This limitation of claim 43 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitation [1L] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [43F] present the promotional content to a user. |

811. For the same reasons discussed above in §V.B.12 for claim limitation [1L], the combination of O'Neill, Jones and Wang discloses this limitation.

## II. GROUND 4: THE COMBINATION OF O'NEILL, JONES AND HOWARD RENDERS OBVIOUS CLAIMS 27-29 AND 44-49

812. The combination of O'Neill, Jones and U.S. Patent No. 6,513,006 ("Howard"; EX-1009) discloses every limitation of claims 27-29 and 44-49. O'Neill, Jones and Howard are analogous art to the '536 patent because all three references are directed to speech recognition systems and software.

813. I have described both O'Neill and Jones above, and I discuss Howard below.

## A. Overview of the Combination

### 1. Howard

814. 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.) Howard discloses the use of "speech recognition and natural language parsing" to "extract the meaning of the user's spoken input" in relation to an electronic activity guide, providing "a fast and far less intrusive interface for a user in selecting a particular activity." (EX-1009, Abstract, 1:66-2:10.)

815. The system in 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 below:

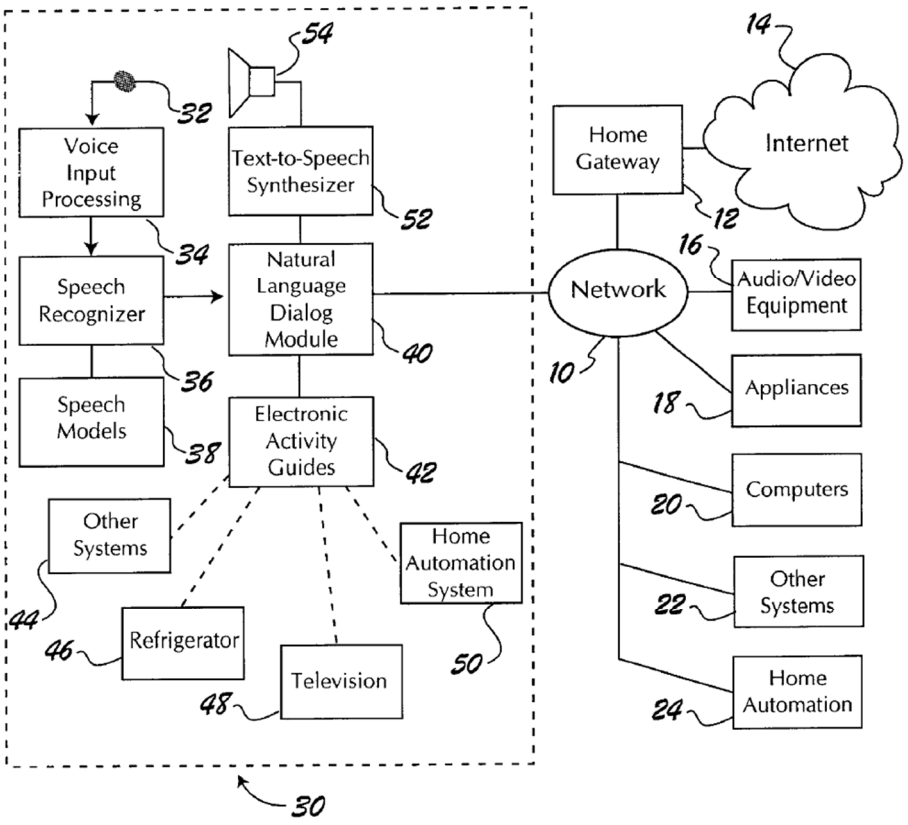



FIG. 1

**Howard, Figure 1**

816. As shown in Figure 1, Howard discloses that "[t]he speech recognizing and natural language processing system 30 of the preferred embodiment includes one of more microphones 32 through which the input speech is provided." (EX-1009, 4:23-25.) "The microphone input is supplied to a voice input processing module 34 which digitizes the input speech and converts it into the form required by the speech recognizer 36." (EX-1009, 4:29-32.) Once the speech is digitized, it 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, 3:57-64.)

817. Howard discloses that each of the devices 16-24 "has independent processing capability and at least some limited data storage ability." (EX-1009, 3:57-60.) "[E]ach device is configured to work independent of its connection to the network." (EX-1009, 3:62-64.) Howard provides a non-exhaustive list of devices that may be used in the system, including "audio/video equipment 16, appliances 18, computers 20 and other systems 22." (EX-1009, 3:37-41.)

**2. Motivation to Combine**

818. 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.

819. A POSITA would have been motivated to make to combine O'Neill, Jones and Howard for numerous reasons.

820. First, as I discussed above in §V.A.3, a POSITA would have been motivated to make such a combination because both O'Neill and Jones suggest that their systems can be combined with other speech processing systems, like Howard.

(EX-1006, [0025] ("The system 10 may perform speech to text conversion using either speech transcription using human transcribers or using conventional speech-to-text processing, also known as automatic speech recognition (ASR)."); EX-1005, [0089] ("The improved dialogue manager will be described by way of example in an accommodation and event booking and payment system. It will be appreciated 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 or other transactional systems.").) Further, 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-1005, [0073] (referring to "domain-specific functionality").)

821. 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.").) A POSITA would thus be motivated to combine the teachings of O'Neill, Jones and Howard because all three systems are attempting to accomplish the same goal: receiving requests, processing requests and responding to requests.

822. Third, adding the targeted advertising of Jones to the teachings of Howard would be intuitive to a POSITA. The Howard system collects various categories of information about users, such as grocery store shopping lists and dinner plans. (EX-1009, 2:37-50, 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.

823. 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).

824. A POSITA would have had a reasonable expectation of success in combining O'Neill, Jones and Howard, 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.

825. 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 (system uses "speech recognition and natural language parsing" to "extract the meaning of the user's spoken input" in relation to an electronic activity guide, providing "a fast and far less intrusive interface for a user in selecting a particular activity.").)

## B. Dependent Claim 27

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

| |
|---|
| [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 |

826. The combination of O'Neill, Jones and Howard discloses "[t]*he 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*."

827. Howard discloses that user requests can be processed by multiple devices associated with the user. As shown in Figure 1 below, the system described in Howard is connected to multiple devices through the network 10.

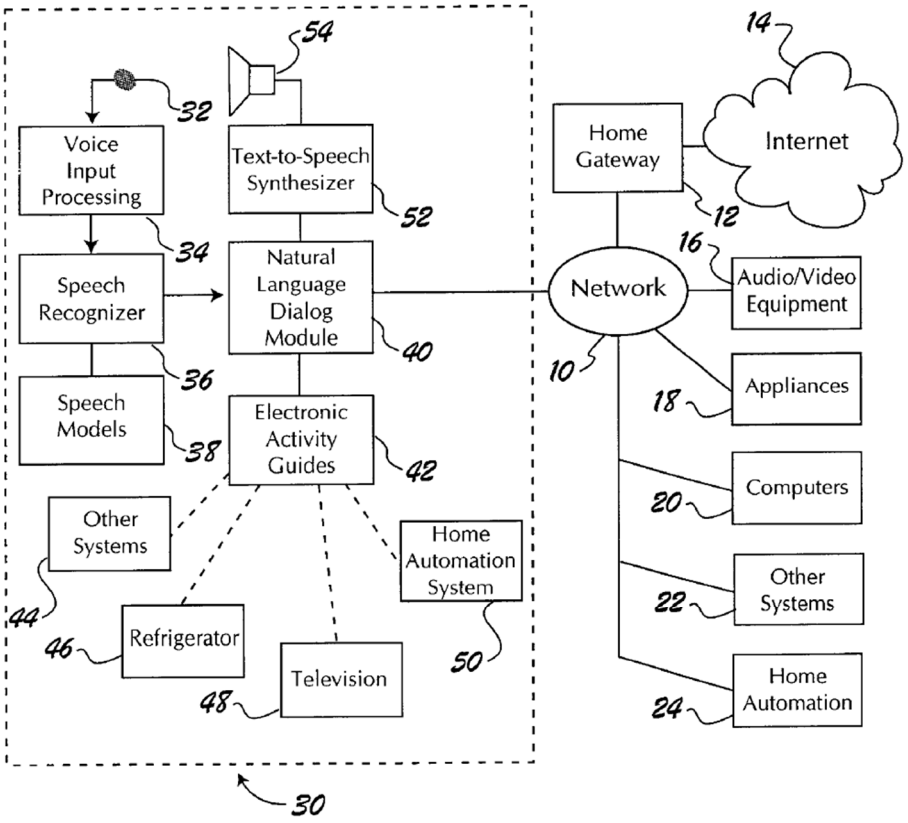



FIG. 1

**Howard, Figure 1**

828. Howard further discloses that the devices connected to the system can process requests in parallel and have "independent processing capability and at least some limited data storage ability." (EX-1009, 3:57-60.)

829. Howard provides an example of a single user utterance which includes multiple requests that are "*processed by a first device associated with the user and … a second device associated with the user*." In the example, "the user, about ready to make a shopping trip, might speak into a microphone associated with the refrigerator and ask, 'What can I make for dinner?' The natural language dialog module would access the appropriate electronic activity guide (the one associated with the refrigerator) and then retrieve information that associates possible dinner menus with items in the refrigerator." (EX-1009, 2:37-43.) The system would then make suggestions about possible menus and what groceries may need to be acquired. (EX-1009, 2:43-45.) 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.)

830. This disclosure from Howard mirrors the disclosure in the specification of the '536 patent related to multiple requests in a single utterance:

> For example, an utterance received in operation 305 may be: "Get me a route to Chang's Restaurant, and call them so I can make a reservation." The utterance may thus include multiple requests, including cross-device requests

> (e.g., calculate a route using a navigation device, and make a telephone call using a mobile phone), as well as cross-application requests (e.g., search for an address and/or phone number using a voice search engine, and calculate a route using a navigation application).

('536 Patent, 50-59.)

831. In this example, a first request asks the system to "calculate a route using a navigation device" which is similar to the request in Howard asking the system to determine what items are not in the user's refrigerator by saying "I think I will make chicken stew tonight." And the second request described in the specification "make a telephone call using a mobile phone" is similar to the request in Howard asking the system to "Please print me a shopping list." In both the '536 patent example and in Howard the requests are handled by different devices: a navigation system and mobile phone in the '536 patent and a refrigerator and printer in Howard.

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

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

832. The combination of O'Neill, Jones and Howard discloses "*wherein the promotional content relates to one or more of the first request or the second request*."

833. This limitation of claim 27 is substantially the same limitation that is addressed in §VII.B.2 regarding claim limitation [25B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [25B] wherein the promotional content relates to one or more of the first request or the second request. | [27B] wherein the promotional content relates to one or more of the first request or the second request. |

834. For the same reasons discussed above in §VII.V.2 for limitation [25B], the combination of O'Neill, Jones and Howard discloses this limitation.

### C. Dependent Claims 28 and 45

[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.

[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.

835. The combination of O'Neill, Jones and Howard discloses "*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.*"

836. 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.)

837. Jones similarly discloses that the system can use multiple devices that are different types, such as "mobile phones, landline phones, specialized voice query terminals, voice-enabled PDA's (personal digital assistants) … touch-tone telephones, two-way radios, wifi phones, cordless phones, portable media devices, or any communication device that allows the user to communicate with the server 16." (EX-1006, [0023].)

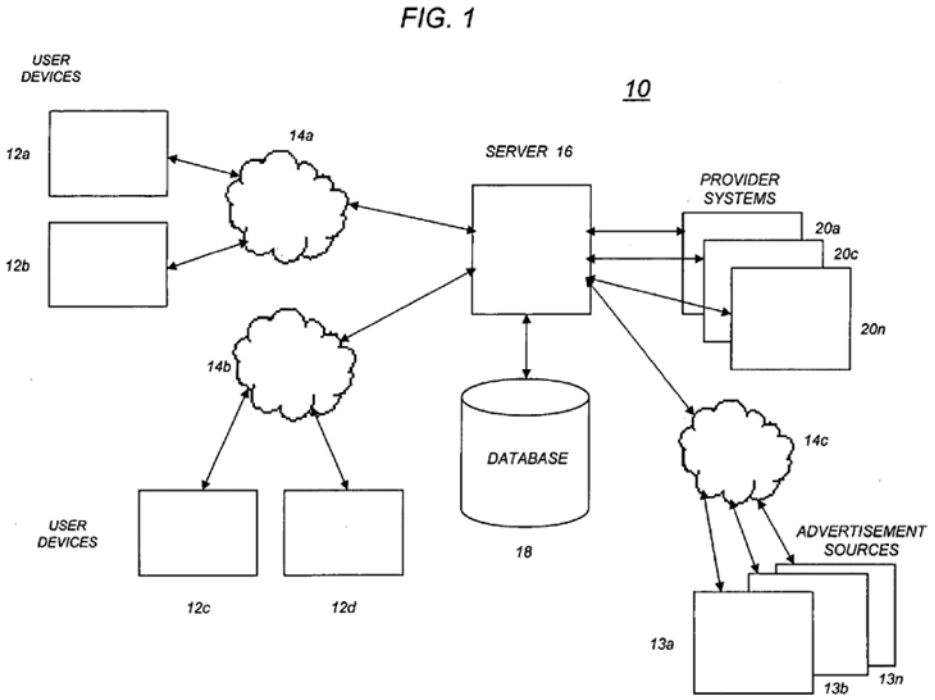
### D. Dependent Claims 29 and 46

> [29] The method of claim 27, wherein the first device and the second device are operably independent of one another.
>
> [46] The method of claim 44, wherein the first device and the second device are operably independent of one another.

838. The combination of O'Neill, Jones and Howard discloses "*wherein the first device and the second device are operably independent of one another*."

839. As discussed above, Howard discloses that each of the devices "*are operably independent of one another*": "each of the devices connected to network 10, such as the devices illustrated at 16-24 in FIG. 1, 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.)

840. Similarly, Jones discloses multiple devices that are "*operably independent of one another*." Figure 1 below demonstrates that user devices 12a, 12b, 12c and 12d can connect independently to the servers 14a and 14b:




**Jones, Figure 1**

## E. Independent Claim 44

### 1. Preamble [44P]

| |
|---|
| [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: |

841. To the extent this preamble is limiting, the combination of O'Neill, Jones and Howard discloses *"[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*."

842. The preamble of claim 44 is substantially the same limitation that is addressed in §V.B.1 regarding the preamble of claims 1 and 32 above. I provide a side-by-side comparison of preambles [1P] and [44P] here:

| | |
|---|---|
| [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: | [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: |

843. For the same reasons discussed above in §V.B.1 for preambles [1P]/[32P], the combination of O'Neill, Jones and Howard discloses this preamble.

### 2. Receiving Natural Language Utterance [44A]

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

844. The combination of O'Neill, Jones and Howard discloses "*receiving, at the one or more physical processors, a first natural language utterance*."

845. This limitation of claim 44 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A]/[32A] above. I provide a side-by-side comparison of limitations [1A] and [44A] here:

| | |
|---|---|
| [1A] receiving, at the one or more physical processors, a first natural language utterance; | [44A] receiving, at the one or more physical processors, a first natural language utterance; |

846. For the same reasons discussed above in §V.B.2 for limitations [1A]/[32A], the combination of O'Neill, Jones and Howard discloses this limitation.

### 3. Providing a Response [44B]

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

847. The combination of O'Neill, Jones and Howard discloses "*providing, by the one or more physical processors, a response to the first natural language utterance*."

848. This limitation of claim 44 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitations [1B]/[32B] above. I provide a side-by-side comparison of limitations [1B] and [44B] here:

| | |
|---|---|
| [1B] providing, by the one or more physical processors, a response to the first natural language utterance; | [44B] providing, by the one or more physical processors, a response to the first natural language utterance; |

849. For the same reasons discussed above in §V.B.3 for limitations [1B]/[32B], the combination of O'Neill, Jones and Howard discloses this limitation.

### 4. Receiving a Second Natural Language Utterance [44C]

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

850. The combination of O'Neill, Jones and Howard discloses "*receiving, at the one or more physical processors, a second natural language utterance relating to the first natural language utterance*."

851. This limitation of claim 44 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitations [1C]/[32C] above. I provide a side-by-side comparison of limitations [1C] and [44C] here:

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

852. For the same reasons discussed above in §V.B.4 for limitations [1C]/[32C], the combination of O'Neill, Jones and Howard discloses this limitation.

### 5. Identifying Requests Associated with the Second Utterance, Wherein First Request is Processed by a First Device and Second Request is Processed by a Second Device [44D]

| |
|---|
| [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; |

853. The combination of O'Neill, Jones and Howard discloses "*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*."

854. This limitation of claim 44 is substantially the same limitation that is addressed in §V.N.1 regarding claim limitations [16A] and [27A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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,<br>[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 | [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; |

855. For the same reasons discussed above in §V.N.1 for limitations [16A] and [27A], the combination of O'Neill, Jones and Howard discloses this limitation.

### 6. Determining Promotional Content Related to One or More Requests [44E]

| |
|---|
| [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 |

856. The combination of O'Neill, Jones and Howard discloses "*determining, by the one or more physical processors, promotional content that relates to one or more of the first request or the second request*."

857. This limitation of claim 44 is substantially the same limitation that is addressed in §VIII.E.5 regarding claim limitation [39E] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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 |

858. For the same reasons discussed above in §VIII.E.5 for limitation [39E], the combination of O'Neill, Jones and Howard discloses this limitation.

### 7. Presenting the Promotional Content [44F]

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

859. The combination of O'Neill, Jones and Howard discloses "*presenting, by the one or more physical processors, the promotional content to the user*."

860. This limitation of claim 44 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitations [1L]/[32L] above. I provide a side-by-side comparison of limitations [1L] and [44F] here:

| | |
|---|---|
| [1L] presenting, by the one or more physical processors, the promotional content to a user. | [44F] presenting, by the one or more physical processors, the promotional content to the user. |

861. For the same reasons discussed above in §V.B.12 for limitations [1L]/[32L], the combination of O'Neill, Jones and Howard discloses this limitation.

## F. Dependent Claim 47

### 1. Receiving Third Natural Language Utterance [47A]

| |
|---|
| [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 |

862. The combination of O'Neill, Jones and Howard discloses "*[t]he 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*."

863. This limitation of claim 47 is substantially the same limitation that is addressed in §V.H.1 regarding claim limitation [7A] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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 | [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 |

864. For the same reasons discussed above in §V.H.1 for limitation [7A], the combination of O'Neill, Jones and Howard discloses this limitation.

### 2. Determining an Interpretation of the Third Natural Language Utterance Based on the Promotional Content [47B]

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

865. The combination of O'Neill, Jones and Howard discloses "*determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content*."

866. This limitation of claim 47 is substantially the same limitation that is addressed in §V.H.2 regarding claim limitation [7B] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [7B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content. | [47B] determining, by the one or more physical processors, an interpretation of the third natural language utterance based on the promotional content. |

867. For the same reasons discussed above in §V.H.2 for limitation [7B], the combination of O'Neill, Jones and Howard discloses this limitation.

### G. Dependent Claim 48

> [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.

868. The combination of O'Neill, Jones and Howard discloses "[t]*he 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*."

869. The elements of claim 48 are substantially the same elements that are addressed in §V.M regarding claim 12 above. I provide a side-by-side comparison of the claims here:

| | |
|---|---|
| [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. | [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. |

870. For the same reasons discussed above in §V.M for claim 12, the combination of O'Neill, Jones and Howard discloses this claim.

## H. Independent Claim 49

### 1. Preamble [49P]

| |
|---|
| [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: |

871. To the extent this preamble is limiting, the combination of O'Neill, Jones and Howard discloses *"[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*."

872. The preamble of claim 49 are substantially the same preambles that is addressed in §V.B.1 regarding claim limitations [1P]/[32P] above. I provide a side-by-side comparison of limitations [32P] and [49P] here:

| | |
|---|---|
| [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: | [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: |

873. For the same reasons discussed above in §V.B.1 for preambles [1P]/[32P], the combination of O'Neill, Jones and Howard discloses this preamble.

### 2. Receive a First Natural Language Utterance [49A]

| |
|---|
| [49A] receive a first natural language utterance; |

874. The combination of O'Neill, Jones and Howard discloses "*receive a first natural language utterance*."

875. This limitation of claim 49 is substantially the same limitation that is addressed in §V.B.2 regarding claim limitation [1A]/[32A] above. I provide a side-by-side comparison of limitations [32A] and [49A] here:

| | |
|---|---|
| [32A] receive a first natural language utterance; | [49A] receive a first natural language utterance; |

876. For the same reasons discussed above in §V.B.2 for limitations [1A]/[32A], the combination of O'Neill, Jones and Howard discloses this limitation.

### 3. Provide Response to First Natural Language Utterance [49B]

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

877. The combination of O'Neill, Jones and Howard discloses "*provide a response to the first natural language utterance*."

878. This limitation of claim 49 is substantially the same limitation that is addressed in §V.B.3 regarding claim limitation [1B]/[32B] above. I provide a side-by-side comparison of limitations [32B] and [49B] here:

| | |
|---|---|
| [32B] provide a response to the first natural language utterance; | [49B] provide a response to the first natural language utterance; |

879. For the same reasons discussed above in §V.B.3 for limitations [1B]/[32B], the combination of O'Neill, Jones and Howard discloses this limitation.

### 4. Receive a Second Natural Language Utterance [49C]

| |
|---|
| [49B] receive a second natural language utterance relating to the first natural language utterance; |

880. The combination of O'Neill, Jones and Howard discloses "*receive a second natural language utterance relating to the first natural language utterance*."

881. This limitation of claim 49 is substantially the same limitation that is addressed in §V.B.4 regarding claim limitations [1C]/[32C] above. I provide a side-by-side comparison of limitations [32C] and [49C] here:

| | |
|---|---|
| [32C] receive a second natural language utterance relating to the first natural language utterance; | [49C] receive a second natural language utterance relating to the first natural language utterance; |

882. For the same reasons discussed above in §V.B.4 for limitations [1C]/[32C], the combination of O'Neill, Jones and Howard discloses this limitation.

### 5. Identify Requests Associated with the Second Utterance, Wherein First Request is Processed by a First Device and Second Request is Processed by a Second Device [49D]

| |
|---|
| [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; |

883. The combination of O'Neill, Jones and Howard discloses "*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*."

884. This limitation of claim 49 is substantially the same limitation that is addressed in §VIII.E.5 regarding claim limitation [44D] above. I provide a side-by-side comparison of the limitations here:

| | |
|---|---|
| [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; | [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; |

885. For the same reasons discussed above in §VIII.E.5 for limitation [44D], the combination of O'Neill, Jones and Howard discloses this limitation.

### 6. Determine Promotional Content Related to One or More Requests [49E]

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

886. The combination of O'Neill, Jones and Howard discloses "*determine promotional content that relates to one or more of the first request or the second request*."

887. This limitation of claim 49 is substantially the same limitation that is addressed in §VII.E.5 regarding claim limitation [39E] above. I provide a side-by-side comparison of the limitations here:

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

888. For the same reasons discussed above in §VII.E.5 for limitation [39E], the combination of O'Neill, Jones and Howard discloses this limitation.

### 7. Presenting the Promotional Content [49F]

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

889. The combination of O'Neill, Jones and Howard discloses "*presenting, by the one or more physical processors, the promotional content to the user*."

890. This limitation of claim 49 is substantially the same limitation that is addressed in §V.B.12 regarding claim limitations [1L]/[32L] above. I provide a side-by-side comparison of limitations [32L] and [49F] here:

| | |
|---|---|
| [32L] present the promotional content to a user. | [49F] present the promotional content to the user. |

891. For the same reasons discussed above in §V.B.12 for limitations [1L]/[32L], the combination of O'Neill, Jones and Howard discloses this limitation.

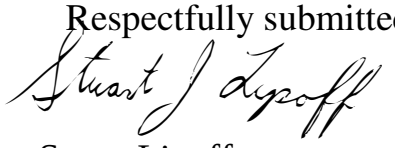
## III. CONCLUSION

892. In signing this Declaration, I recognize that the Declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in this proceeding. If required, I will appear for cross-examination at the appropriate time. I reserve the right to offer opinions relevant to the invalidity of the challenged claims at issue and/or offer testimony in support of this Declaration.

893. I hereby declare that all statements made herein of my own knowledge are true and that all statements are made on information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001.

Dated: April 21, 2025

Respectfully submitted,



Stuart Lipoff

## APPENDIX

| Exh. | Reference |
|---|---|
| 1001 | U.S. Patent 8,886,536 to Freeman & Kennewick ("the '536 patent") |
| 1002 | Prosecution History for U.S. Patent 8,886,536 |
| 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 *VB Assets, LLC v. Amazon.com Servs. LLC,* No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019) ("Amazon Litigation Mem. Op. Regarding Claim Construction") |
| 1011 | Second Amended Joint Claim Construction Chart from *VB Assets, LLC v. Amazon.com Servs. LLC,* No. 1:19-cv-01410-MN (D. Del., filed July 29, 2019) ("Amazon Litigation Second Am. Joint Claim Construction Chart") |
| 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., *Designing Interactive Speech Systems: From First Ideas to User Testing*, Springer (1998) ("Bernsen") |
| 1015 | Xuedong Huang et al., *Spoken Language Processing*, Prentice Hall PTR (2001) ("Huang") |
| 1016 | Michael F. McTear, *Spoken Dialogue Technology* (2004) ("McTear") |

| Exh. | Reference |
|---|---|
| 1017 | Andrew Kehler, *The Effect of Establishing Coherence in Ellipsis and Anaphora Resolution*, 31st Annual Meeting of the Association for Computational Linguistics (1993) ("Kehler") |
| 1018 | Sebastian Möller, *Quality of Telephone-Based Spoken Dialogue Systems*, Springer (2005) ("Möller") |
| 1019 | Phillip R. Cohen et al., *An Open Agent Architecture*, Proceedings of the 1994 AAAI Spring Symposium, Software Agents (1994) ("Cohen") |
| 1020 | Stephanie Seneff et al., *Multimodal Discourse Modelling in a Multi-User Multi-Domain Environment*, 4th International Conference on Spoken Language Processing (Oct. 1996) |
| 1021 | Stephanie Seneff et al., *GALAXY-II A Reference Architecture for Conversational System Development*, Proceedings of the 5th International Conference on Spoken Language Processing (Nov. 1998) ("GALAXY-II") |
| 1022 | Ian O'Neill et al., *Cross Domain Dialogue Modelling: An Object-Based Approach*, Proceedings of the 8th International Conference on Spoken Language Processing (2004) ("O'Neill (2004)") |
| 1023 | Botond Pakucs, *SesaME: A Framework for Personalised and Adaptive Speech Interfaces*, Proceedings of the 2003 EACL Workshop on Dialogue Systems: Interaction, Adaptation, and Styles of Management (2003) ("Pakucs") |
| 1024 | Mikio Nakano et al., *Grounding New Words on the Physical World in Multi-Domain Human-Robot Dialogues*, Papers from the 2010 AAAI Fall Symposium (2010) ("Nakano") |
| 1025 | Honghua (Kathy) Dai et al., *Detecting Online Commercial Intention (OCI)*, Proceedings of the 15th International Conference on World Wide Web (2006) ("Dai") |
| 1026 | Interactive Advertising Bureau, *IAB Internet Advertising Revenue Report* (May 2007) ("Internet Advertising Bureau Report") |

| Exh. | Reference |
|---|---|
| 1027 | Joe Plummer et al., *The Online Advertising Playbook*, (2007) ("Plummer") |
| 1028 | Ganesh Iyer et al., *The Targeting of Advertising*, INSEAD Working Paper Series (Feb. 2004) ("Iyer") |
| 1029 | Alex Kinnier, *Why We're Buying DoubleClick*, Google Blog (June 26, 2007), https://googleblog.blogspot.com/2007/06/why-were-buying-doubleclick.html |
| 1030 | Daniel C. Fain & Jan O. Pedersen, *Sponsored Search: A Brief History*, 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, *Internet Advertising Via "Behavioral Targeting"*, USAToday.com (Sept. 28, 2004) ("Duane Morris") |
| 1033 | Kate Burns, *Ten Golden Rules to Search Advertising*, 6(3) Interactive Marketing 248 (2005) ("Burns") |