

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

IPR2025-00868

U.S. Patent No. 8,515,765

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,515,765

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

TABLE OF CONTENTS

I.	Grounds for Standing.....	1
II.	Identification of Challenge	1
	A. Prior Art.....	1
	B. Grounds For Challenge	2
III.	The '765 Patent.....	2
	A. Patent Overview	2
	B. Technical Background.....	7
	C. Level of Ordinary Skill	10
	D. Claim Construction.....	10
IV.	GROUND 1: Smith & Hipp-Guinn Combination Renders Obvious Claims 1-14.....	10
	A. Combination Overview	10
	1. Smith & Hipp	10
	2. Guinn.....	12
	3. Motivation to Combine	14
	B. Independent Claims 1, 10.....	16
	1. Preambles [1P]/[10P]	16
	2. Receiving Natural Language Utterance [1A]/[10A].....	18
	3. Determining the Conversation Type Corresponding to the Natural Language Utterance [1B]/[10B]	22
	4. Generating a Response to the Natural Language Utterance [1C]/[10C]	30

C.	Claims 2 and 11 – “format of the response is adapted to provide discrete piece of information, ... a plurality of pieces of information, or obtain information.”	33
1.	Determining a Conversational Goal Corresponding to the Natural Language Utterance [2A]/[11A]	34
2.	Format of the Response Is Adapted to Provide ... [2B]/[11B]	36
D.	Claim 3 – “discrete” responses shorter than “plurality of pieces” responses	41
E.	Claims 4 and 12 – “determining conversation type” based on “conversational goal.”	42
1.	Determining a Conversational Goal Corresponding to the Natural Language Utterance [4A]/[12A]	43
2.	Conversation Type Is Determined to Be a Query, Didactic, or Exploratory Conversation [4B]/[12B]	43
F.	Claims 5 and 13 – “conversation type” determined based on “role”	45
G.	Claims 6, 7, and 14 – “conversation type” determined based on “allocation of information.”	50
H.	Claim 8 - “only provides information requested ... limits the user’s future input”	57
I.	Claim 9 - “user controls the conversation or ... provides information”	58
V.	Discretionary Denial is Not Appropriate.....	63
VI.	Mandatory Notices.....	64
A.	Real Party In Interest.....	64
B.	Related Matters.....	64
C.	Notice of Counsel and Service Information.....	64

VII. Conclusion	65
-----------------------	----

EXHIBIT LIST

Exh.	Reference
1001	U.S. Patent 8,515,765 to Baldwin et al.
1002	Prosecution History for U.S. Patent 8,515,765
1003	Declaration of Stuart Lipoff in Support of Inter Partes Review of U.S. Patent 8,515,765
1004	Curriculum Vitae of Stuart Lipoff
1005	Ronnie. W. Smith, et al., <i>An Architecture for Voice Dialog Systems Based on Prolog-Style Theorem Proving</i> , 21(3) Computational Linguistics 281 (1995) (“Smith & Hipp”)
1006	Curry I. Guinn, <i>Mechanisms for Mixed-Initiative Human-Computer Collaborative Discourse</i> , Proceedings of the 34th Annual Meeting of the Association for Computational (1996) (“Guinn”)
1007	Declaration of June Ann Munford in Support of Inter Partes Review of U.S. Patent 8,515,765
1008	Sebastian Möller, <i>Quality of Telephone-Based Spoken Dialogue Systems</i> , Springer (2005) (“Möller”)
1009	Michael F. McTear, <i>Spoken Dialogue Technology</i> (2004) (“McTear”)
1010	Xuedong Huang et al., <i>Spoken Language Processing</i> , Prentice Hall PTR (2001) (“Huang”)
1011	Peter Wiemer-Hastings et al., <i>RMT: A Dialog-Based Research Methods Tutor With or Without a Head</i> , 7th International Conference on Intelligent Tutoring Systems, Vol. 3220 (Jan. 2004)
1012	Arthur C. Graesser et al., <i>AutoTutor: A Tutor with Dialogue in Natural Language</i> , 36(2) Behavior Research Methods, Instruments, & Computers 180 (2004)
1013	Niels Ole Bernsen et al., <i>Designing Interactive Speech Systems: From First Ideas to User Testing</i> , Springer (1998) (“Bernsen”)

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

(“Petitioners”) petition for *inter partes* review of claims 1-14 of U.S. Patent 8,515,765 (“the ’765 patent”; EX-1001).

I. Grounds for Standing

Petitioners certify the ’765 patent is available for *inter partes* review; Petitioners are not barred or estopped.

II. Identification of Challenge

A. Prior Art

The ’765 patent, filed on October 3, 2011, claims priority through continuation to U.S. Patent 8,073,681, filed October 16, 2006. Petitioners do not acquiesce the ’765 patent is entitled to priority of the listed priority patent. Regardless, each applied reference was filed or published before October 16, 2006.

1. **R.W. Smith, et al., “An Architecture for Voice Dialog Systems Based on Prolog-Style Theorem Proving” (1995)** (“Smith & Hipp”; EX-1005), published September 1995, is prior art under 35 U.S.C. §102(b). (*See* Declaration of June Munford; EX-1007, ¶¶13-16.)

2. **Curry I. Guinn, “Mechanisms for Mixed-Initiative Human-Computer Collaborative Discourse” (1996)** (“Guinn”; EX-1006), published June 1996, is prior art under 35 U.S.C. §102(b). (*See* EX-1007, ¶¶17-20.)

B. Grounds For Challenge

Ground		Claims	Prior Art
1	103	1-14	Smith & Hipp and Guinn

III. The '765 Patent

A. Patent Overview

The '765 patent discloses three structural components to the claimed method and system: (1) a “user”, (2) a “voice-enabled device”, and (3) “a speech engine.” (Declaration of Stuart Lipoff; EX-1003, ¶78.) The disclosures corresponding to the claimed “voice-enabled device” and “speech engine” are contained in Figure 1 (below), which is described as an “exemplary block diagram of a system architecture.” (EX-1003, ¶79.).

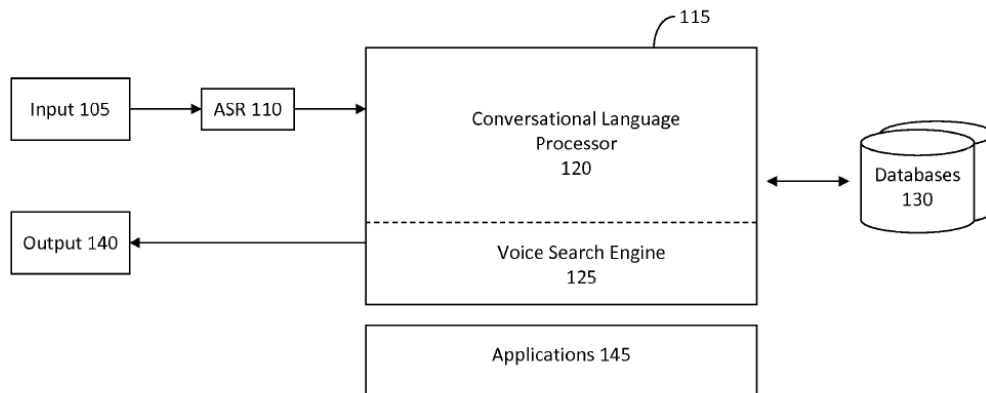


Figure 1

'765 Patent, Figure 1

Per this disclosed architecture, the “utterance component of input 105 may be processed by a speech recognition engine 110 ... to generate one or more preliminary interpretations of the utterance. (EX-1001, 7:31-35.) “The one or more preliminary interpretations generated by the speech recognition engine 110 may then be provided to a conversational speech engine 115 for further processing.” (EX-1001, 7:44-47.)

Most relevant to the claims at issue here is the so-called “cooperative conversational model” for a “Cooperative Conversational Voice User Interface” as depicted in Figure 3 (below). (EX-1003, ¶81.)

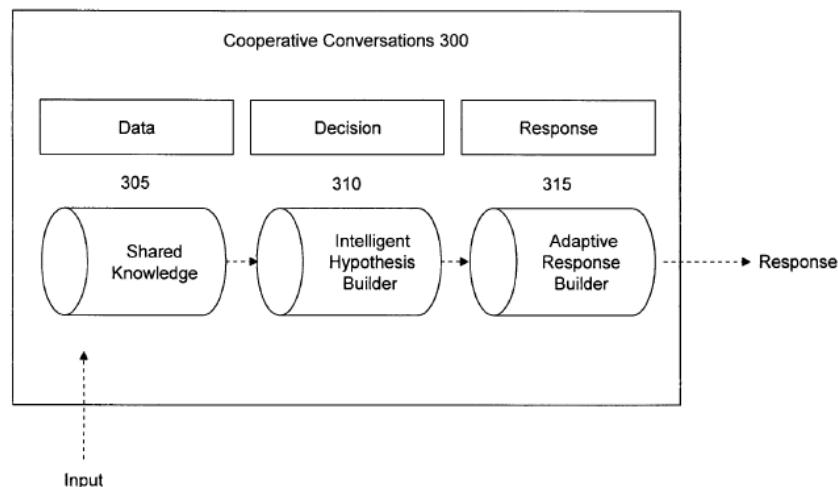


FIG. 3

'765 Patent, Figure 3

The specification provides that “a cooperative conversation may take incoming data (shared knowledge) 305 to inform a decision (intelligent hypothesis builder) 310, and then may refine the decision and generate a response (adaptive response

builder) 315.” (EX-1001, 13:34-38.). The adaptive response may vary tone, grammar, or formality, and “may frame responses to influence a user reply utterance” (EX-1001, 18:18-43.)

Relevant here, the ’765 patent discloses that one role of the Intelligent Hypothesis Builder is to “generate a hypothesis as to a conversation type.” (EX-1001, 14:43-44). The ’765 patent discloses that the Intelligent Hypothesis Builder can consider “conversational goals, participant roles, and/or an allocation of information among the participants.” (EX-1001, 14:45-46.)

“Conversational goals may broadly include: (1) getting a discrete piece of information or performing a discrete task, (2) gathering related pieces of information to make a decision, and/or (3) disseminating or gathering large amounts of information to build expertise.” (EX-1001, 14:46-50.)

“Participant roles may broadly include: (1) a leader that controls a conversation, (2) a supporter that follows the leader and provides input as requested, and/or (3) a consumer that uses information.” (EX-1001, 14:51-54.)

With respect to the allocation of information, the ’765 patent discloses that “[i]nformation may be held by one or more of the participants at the outset of a conversation, where a participant may hold most (or all) of the information, little (or none) of the information, or the information may be allocated roughly equally amongst the participants.” (EX-1001, 14:54-58.)

The '765 patent discloses three types of conversations that may be identified by the Intelligent Hypothesis Builder: Query, Didactic, and Exploratory. (*See generally* EX-1001, 15:1-40.) The patent includes the following disclosures related to these three “types” of conversations:

- **Query:** “[A] query conversation may include a conversational goal of getting a discrete piece of information or performing a particular task, where a leader of the query conversation may have a specific goal in mind and may lead the conversation toward achieving the goal. The other participant may hold the information and may support the leader by providing the information.” (EX-1001, 15:1-7.)
- **Didactic:** “In a didactic conversation, a leader of the conversation may control information desired by a supporter of the conversation. The supporter’s role may be limited to regulating an overall progression of the conversation and interjecting queries for clarification.” (EX-1001, 15:7-12.)
- **Exploratory:** “In an exploratory conversation, both participants share leader and supporter roles, and the conversation may have no specific goal, or the goal may be improvised as the conversation progresses.” (EX-1001, 15:11-16.)

The '765 patent further discloses that the “Intelligent Hypothesis Builder 310 may broadly categorize a conversation (or utterance) according to the following diagram:”

	Participant A User	Participant B Voice User Interface
QUERY		
GOAL	Get information/action	Provide information/action
ROLE	Leader/Consumer	Supporter/Dispenser
INFORMATION ALLOCATION	Less	More
DIDACTIC		
GOAL	Get information	Provide information
ROLE	Follower/Consumer	Leader/Dispenser
INFORMATION ALLOCATION	Less	More
EXPLORATORY		
GOAL	Gather/share information	Gather/share information
ROLE	Follower/Consumer and Leader/Dispenser	Follower/Consumer and Leader/Dispenser
INFORMATION ALLOCATION	Equal or alternating	Equal or alternating

(EX-1001, 15:20-40.) The specification provides that “a conversation type may change dynamically as information is reallocated from one participant to another, or as conversational goals change based on the reallocation of information.” (EX-1001, 14:64-67.)

Outside of these disclosures, the ’765 patent does not teach *how* the disclosed Intelligent Hypothesis Builder determines a conversation type from an utterance or a conversation. (EX-1003, ¶90.) Nor does it explain how the Intelligent Hypothesis Builder determines whether the conversational goal is to get information or provide information; whether the user role is one of follower,

consumer, leader, supporter, or dispenser; nor the allocation of information between the participants. (EX-1003, ¶90.)

B. Technical Background

By the earliest possible priority date of the '765 patent, spoken dialog systems were well known and commercially available in various fields, including directory assistance, customer service, dictation, translation, and tutorial systems. Spoken Dialog Systems include various componentry, including speech recognizers, semantic analyzers, dialog managers, response generators, and speech synthesizers. (EX-1003, ¶¶34-42.) However, the claims of the '765 patent relate principally to the function of the dialog manager, which is sometimes called the “central component” of any spoken dialog system. (EX-1003, ¶43.)

Dialog managers take on multiple roles. (EX-1003, ¶46.) Relevant here, dialog managers are responsible for controlling the related concepts of dialog initiative and dialog flow. (EX-1003, ¶¶47-48.) Dialog initiative refers to the concept of control; that is, who is driving the conversation forward towards an objective or goal. (EX-1003, ¶¶49-50.) Broadly speaking, some spoken dialog systems are set up to be system-led (the system always has the initiative) and some are set up to be user-led (where the user always has the initiative). (EX-1003, ¶49.) In addition, it is possible for systems to be “mixed-initiative”, which means that either the system or the user can be in control and the choice of the next step is

determined dynamically, based on the current state of the dialog. (EX-1003, ¶¶49-50.) These systems are more advanced as they require the dialog manager to act as an “intelligent agent” that is able to make rational decisions about how to manage the dialog. (EX-1003, ¶50.) However, by the time of the earliest priority date of the ’765 patent, multiple such systems had been developed around the world. (EX-1003, ¶75.)

By the earliest possible priority date of the ’765 patent, many different types of spoken dialog systems were in existence. (EX-1003, ¶¶58-75.) While there are no hard and fast taxonomies for the different types of spoken dialog systems, experts have often categorized systems based on how initiative is distributed between the system and the user, the level of cooperativity between system and user, and the observable behavior of the system. (EX-1003, ¶¶58-60.)

Along these dimensions, by 2006, three broad categories of spoken dialogue systems were prominent: (1) information retrieval / task-oriented dialog systems, (2) instructional or tutorial dialog systems, and (3) conversational dialog systems. (EX-1003, ¶60.)

Information retrieval systems were most common and most familiar spoken dialogue systems as of the earliest priority date of the ’765 patent. (EX-1003, ¶¶61-66.) In these systems, the user would have dialog initiative and would lodge discrete requests for information from the system. (EX-1003, ¶¶62-63.) Petitioner’s

expert, Stuart Lipoff, provides an example of the Air Traffic Information System projects, (EX-1003, ¶64), but the reader may be more familiar with services like Moviefone (*i.e.*, 777-FILM). These systems were structured to provide the information requested by the user, and little more. (EX-1003, ¶¶61-65.)

Instructional or tutorial dialog systems, by contrast, placed the control of the dialog in the hands of the system. (EX-1003, ¶¶67-72.) The objective of these systems was to convey knowledge from the system to the user. (EX-1003, ¶67.) These systems were sometimes used in the context of teaching, onboarding, or walk throughs (EX-1003, ¶67.) and it was the goal of the system to ensure that learning objectives were met. (EX-1003, ¶68.) A challenge for such systems was ensuring that the system was able to maintain dialog control. (EX-1003, ¶70.) This is because, if the user were permitted to have dialog control they may provide lengthy and complex utterances that the system had no ability to understand or respond to. (EX-1003, ¶70.) Various techniques were developed to support the system maintaining dialog control, including through phrasing the system's utterances in ways that did not relinquish dialog control. (EX-1003, ¶¶71-72.)

Conversational dialog systems are the most advanced and capable type of spoken dialog systems. (EX-1003, ¶¶73-74.) This configuration provides for a back-and-forth in who possesses the initiative in the system. (EX-1003, ¶¶73-74.) The system also allows for less defined goals or motivations, or more flexibility,

and attempts to understand the motivation behind the dialog as the dialog progresses. (EX-1003, ¶¶73-74.) While the technical complexity of these systems meant that they were last on the scene, by the earliest priority date of the '765 patent, a number of such systems had in fact been developed. (EX-1003, ¶75.)

C. Level of Ordinary Skill

A person of ordinary skill in the art (“POSITA”) would have a bachelor’s degree in computer science, computer engineering, electrical engineering, or related field in computing technology, and two years of experience with automatic speech recognition and natural language understanding, or equivalent education, research experience, or knowledge. (EX-1003, ¶¶99-100.)

D. Claim Construction

The Board need not expressly construe any term for purposes of this proceeding because Petitioners establish the challenged claims are unpatentable using a construction of the claims’ plain meaning. *See Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999).

IV. GROUND 1: Smith & Hipp-Guinn Combination Renders Obvious Claims 1-14.

A. Combination Overview

1. Smith & Hipp

Smith & Hipp discloses a voice dialog system with capabilities such as “problem solving, natural language input and output, a user model, variable

initiative, and the use of expectation for speech error correction and plan recognition.” (EX-1005, 281.) The system is designed to support real-time dialog in task domains such as electronic circuit repair and is capable of generating responses, managing user input, switching between dialog threads, and modeling user knowledge. (EX-1003, ¶104.)

Smith & Hipp discloses that its system may operate using four different levels (or modes) of dialog initiative, covering a spectrum ranging from the computer having complete control to the user having complete control: (1) directive; (2) suggestive; (3) declarative; and (4) passive. (EX-1005, 282-83; EX-1003, ¶¶107-110.)

In **directive** mode, the computer is the leader and maintains control of the conversation at each step: “The computer has complete dialog control. It recommends a subgoal for completion and will use whatever dialog is necessary to obtain the needed item of knowledge related to the subgoal.” (EX-1005, 282.)

On the other end of the control spectrum is **passive** mode, where the *user* is the leader and controls the conversation: “The user has complete dialog control. The computer responds directly to user questions and passively acknowledges user statements without recommending a subgoal as the next course of action.” (EX-1005, 283.)

Suggestive and **declarative** provide “intermediate” levels of control between the user and the computer. (EX-1005, 289; EX-1003, ¶111.) In **suggestive** mode, the system still predominantly controls the dialog but in a less forceful manner—the system may propose the next subgoal or action rather than command it, and the user is allowed to make minor digressions or clarifications related to the task at hand. (EX-1005, 283, 289.) In **declarative** mode, the user generally controls the conversation, but the computer remains cooperative by adding relevant information or reminders unprompted, if they might help the user. (EX-1005, 283, 289.)

While Smith & Hipp discloses the use of variable initiative modes in a conversation as well as theoretically how variable initiative (switching between initiative modes) can be achieved by the system in an ongoing conversation (*see, e.g.,* EX-1005, 289-90), it notes that the system can still be improved in the future with an automatic mechanism for switching initiative modes: “A number of important unsolved problems need attention to enable further progress. For example, an automatic mechanism is needed for setting the appropriate level of initiative as an interaction proceeds.” (EX-1005, 316.)

2. Guinn

Guinn “examine[s] mechanisms for automatic dialogue initiative setting.” (EX-1006, 278.) In particular, Guinn states, “We show how to incorporate

initiative changing in a task-oriented human-computer dialogue system, and we evaluate the effects of initiative both analytically and via computer-computer dialogue simulation.” (EX-1006, 278.) Like Smith & Hipp, Guinn recognized that Smith & Hipp’s system did not include automatic initiative switching and sought to provide that feature: “The dialogue model of Smith (1995) allows for either the computer or the user to assume degrees of initiative; however, Smith presents no algorithm for the computer to change initiative during a dialogue. Our model of mixed-initiative dialogue allows either participant to be in control of the dialogue at any point in time.” (EX-1006, 278.)

In Guinn’s model for mixed-initiative dialogue, “initiative levels for each goal are defined during the interaction based on 1) explicit and implicit initiative-changing utterances and 2) competency evaluation.” (EX-1006, 279.) Applying these two factors, Guinn investigated four initiative-setting schemes:

- “**Random**” mode, where “one agent is given initiative at random in the event of a conflict.” (EX-1006, 281.)
- “**SingleSelection**” mode, where “the more knowledgeable agent (defined by which agent has the greater total percentage of knowledge) is given initiative ... and initiative is set throughout the dialogue.” (EX-1006, 281.)
- “**Continuous**” mode, where “the more knowledgeable agent (defined by which agent’s first-ranked branch is more likely to succeed) is initially given initiative. If that branch fails, this agent’s second-ranked branch is

compared to the other agent's first-ranked branch with the winner gaining initiative." (EX-1006, 281.)

- **"Oracle"** mode, where "an all-knowing mediator selects the agent that has the correct branch ranked highest in its list of branches." (EX-1006, 281)

(EX-1003, ¶¶127-28.)

Guinn's study on automatic initiative switching concluded that "SingleSelection and Continuous modes perform significantly better than Random mode. On average Continuous mode results in 40% less branches searched per goal than Random" and "Continuous mode performs between 15-20% better than SingleSelection." (EX-1006, 281-82; EX-1003, ¶129.)

This Petition focuses on Guinn's Continuous mode.

3. Motivation to Combine

A POSITA would have been motivated to integrate Guinn's automatic mechanism for initiative switching and selection, particularly in Continuous mode, with Smith & Hipp's voice dialog system for a number of reasons. (EX-1003, ¶¶132-134.)

First, Smith & Hipp expressly motivates the combination by identifying an "important unsolved problem[]" in its system, which is achieving the described automatically switching between initiative modes: "A number of important unsolved problems need attention to enable further progress. For example, **an**

automatic mechanism is needed for setting the appropriate level of initiative as an interaction proceeds.” (EX-1005, 316; EX-1003, ¶135.) A POSITA would therefore be motivated look to other references, like Guinn, to address this important unsolved problem in Smith & Hipp’s system. (EX-1003, ¶135.)

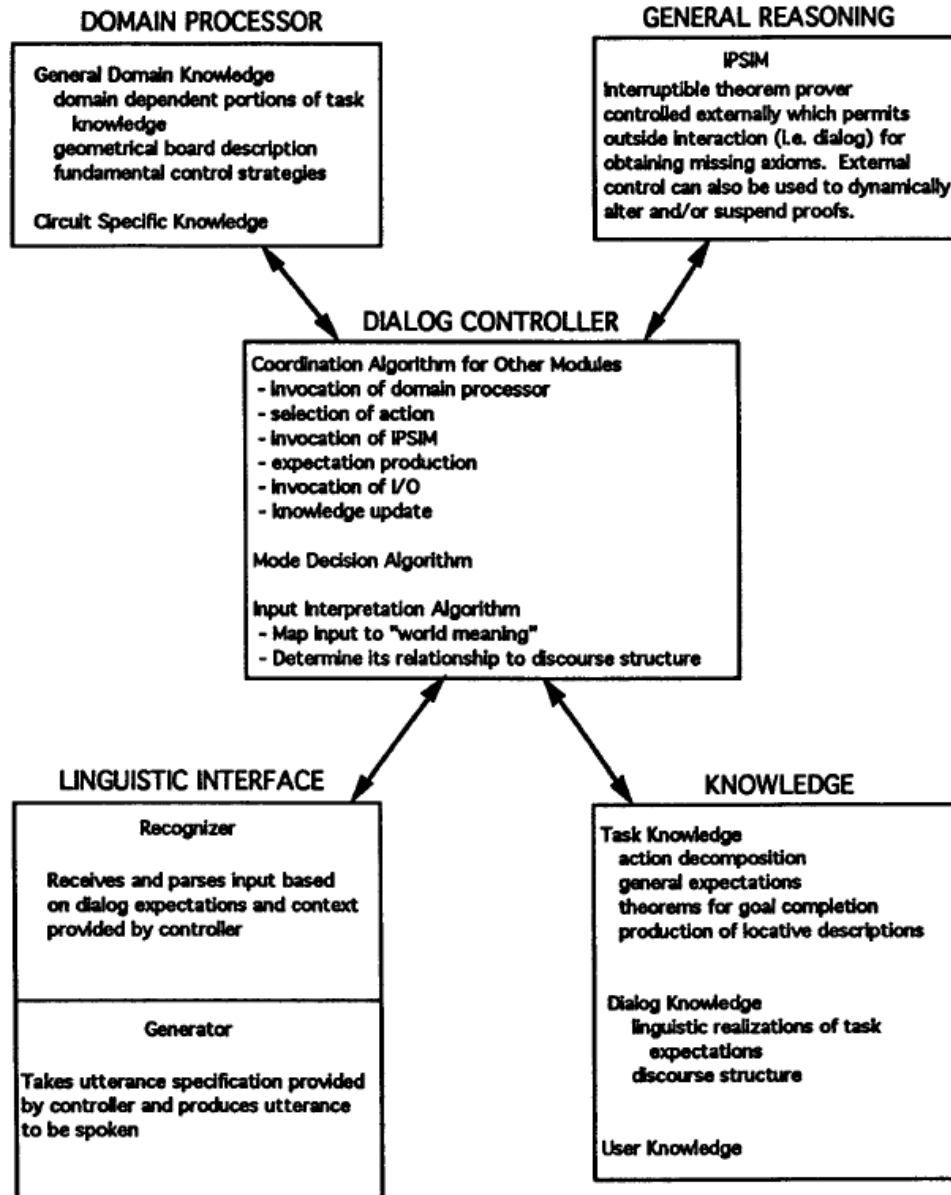
Second, Guinn also recognizes this motivation, noting: “The dialogue model of Smith (1995) allows for either the computer or the user to assume degrees of initiative; however, **Smith presents no algorithm for the computer to change initiative during a dialogue.**” (EX-1006, 278.) Given this, a POSITA would understand Guinn’s teachings to be directly applicable to the system disclosed in Smith & Hipp. (EX-1003, ¶136.) Notably, Guinn confirms that its teachings resolve the implementation problem identified by Smith & Hipp: “Our model of mixed-initiative dialogue allows either participant to be in control of the dialogue at any point in time.” (EX-1006, 278.)

The results of the combination would have been predictable, and a POSITA would have had a reasonable expectation of success, because Smith & Hipp recognizes the need for improvement by adding an automatic initiative switching mechanism, and Guinn’s research was expressly intended to improve the Smith & Hipp system. (EX-1005, 316; EX-1006, 278; EX-1003, ¶¶131-143.)

B. Independent Claims 1, 10

1. Preambles [1P]/[10P]

The Smith & Hipp-Guinn combination discloses a “*system*” that practices a “*method*” of “*providing a voice interface.*” (EX-1003, ¶¶145-49.) Smith & Hipp provides “[a] pragmatic architecture for voice dialog machines.” (EX-1005, 281.) The architecture of the Smith & Hipp system is depicted in Figure 2 below, “where five major subsystems are shown: the dialog controller, the domain processor, the knowledge base, the general reasoning system, and the linguistic interface.” (EX-1005, 297.)



Smith & Hipp, Figure 2

The high-level function of each subsystem is as follows:

- The “**dialog controller**” is “the overall ‘boss’ of the dialog processing system. It formulates goals at the top level to be passed on to the theorem-proving stage. It determines the role that the computer plays in the dialog by determining how user inputs relate to already established

dialog as well as determining the type of response given.” (EX-1005, 297.)

- The “**domain processor**” “contains much of the information about the application domain. It receives from the controller a request for the next suggested goal to be undertaken, and it returns to the controller its suggestion along with expected results from attempting the test.” (EX-1005, 298.)
- The “**reasoning system**” “receives as input goals to be proven and commands to start, stop, and furnish information. It yields as output theorems that are proven and status reports on the proof in progress.” (EX-1005, 300.)
- The “**knowledge base**” “is the repository of information about task-oriented dialogs.” (EX-1005, 300.)
- The “**linguistic interface**” “receives the spoken inputs from the user and returns spoken outputs.” (EX-1005, 300.)

These subsystems together implement the voice dialog system of Smith & Hipp. (EX-1003, ¶146-148.) Accordingly, the Smith & Hipp-Guinn combination discloses both a “*method*” and a “*system*” for “*providing a voice interface.*” (EX-1003, ¶149.)

2. Receiving Natural Language Utterance [1A]/[10A]

The Smith & Hipp-Guinn combination discloses “*receiving, at a speech engine, a natural language utterance from a voice-enabled device, the natural language utterance corresponding to a conversation type*” [1A] and “*a speech engine configured to receive a natural language utterance from a voice-enabled*

device, the natural language utterance corresponding to a conversation type”

[10A].

The Smith & Hipp system is a “*speech engine*,” consistent with the disclosures of the ’765 patent. (EX-1003, ¶¶152-53.) For example, the ’765 patent discloses that, as shown in Figure 1 below, “[c]onversational speech engine 115 may communicate with one or more databases 130 to generate an adaptive conversational response, which may be returned to the user as an output 140.” (EX-1001, 7:50-53.)

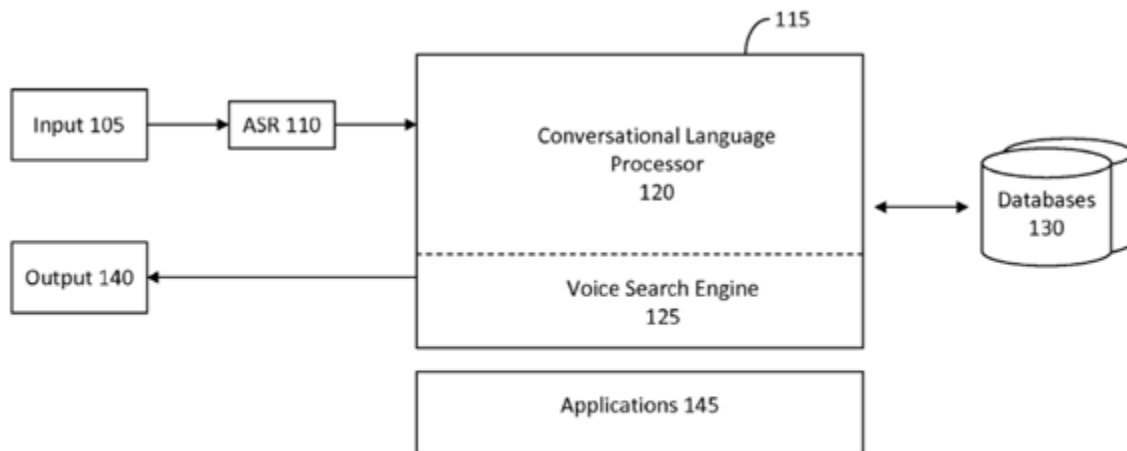
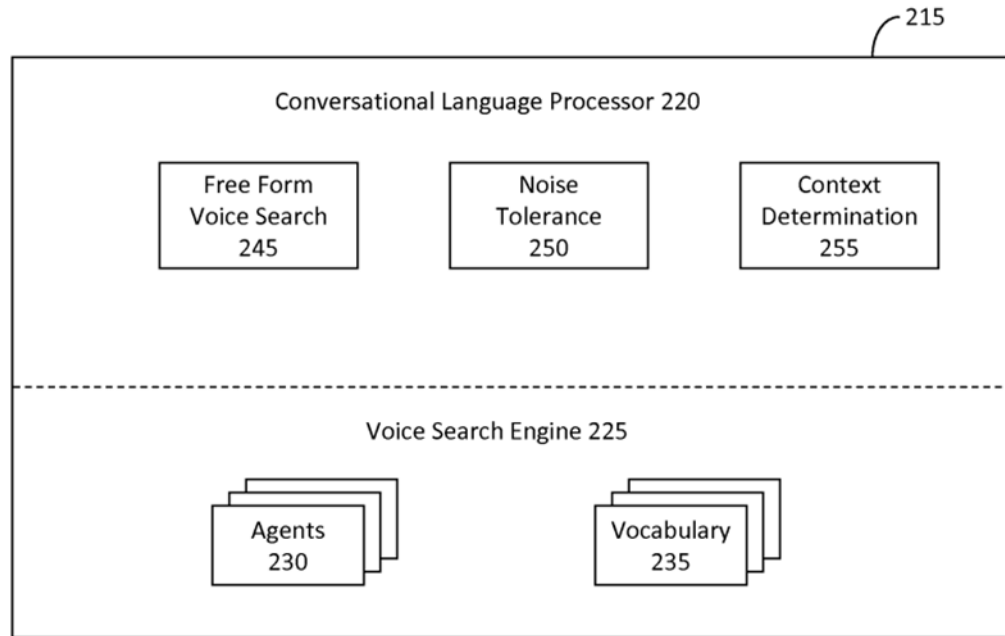


Figure 1

’765 patent, Figure 1

The ’765 patent further teaches that, as shown in Figure 2 below, a conversational speech engine “may include a conversational language processor 220 that generates an adaptive conversational response to a request or series of

requests using a free form voice search module 245, a noise tolerance module 250, and/or a context determination process 255.” (EX-1001, 8:3-8.)



'765 patent, Figure 2

A POSITA would understand that the various components of the '765 patent's system shown in Figures 1 and 2, including the speech engine, are implemented in software on one or more devices, including, for example, a user device like a laptop computer. (EX-1003, ¶¶152-53; EX-1001, 7:61-67 (“[I]n various implementations, the speech recognition engine 110, conversational speech engine 115, and/or databases 130 may reside locally (e.g., on a user device), remotely (e.g., on a server), or a hybrid model of local and remote processing may be used (e.g., lightweight applications may be processed locally while computationally intensive applications may be processed remotely).”).)

The Smith & Hipp system includes software modules to perform the same functions as the “speech engine” disclosed in the ’765 patent—namely, the system generates an adaptative conversational response to voice inputs received from a user. (EX-1003, ¶156; *see, e.g.*, EX-1005, 297 (“[The dialog controller] determines the role that the computer plays in the dialog by determining how user inputs relate to already established dialog as well as determining the type of response given.”).) Smith & Hipp explains the breakdown of the software for the voice dialog system as follows:

The dialog system exclusive of the parser and error correction code consists of about 17,000 lines of Prolog (and this includes some comments) apportioned as follows: Dialog Controller procedural mechanisms (including IPSIM) 15%, Dialog Controller knowledge base 11%, Domain Processing procedural mechanisms 25%, Domain Processing knowledge base 14%, Linguistic Interface including much language generation code 30%, miscellaneous 5%.

(EX-1005, 303.)

The Smith & Hipp system (“speech engine”) “*receiv[es] a natural language utterance from a voice-enabled device, the natural language utterance corresponding to a conversation type.*” Smith & Hipp’s system includes a linguistic interface that receives voice inputs from users, also referred to as a “speech utterance” in Smith & Hipp. Specifically, Smith & Hipp teaches that the linguistic interface “**receives the spoken inputs from the user** and returns spoken

outputs.... The inputs to the voice input system for each utterance are the set of expectations from the dialog controller and the speech utterance from the user. The output from this system is a GADL [Goal and Action Description Language] meaning representation.” (EX-1005, 300; *see also* EX-1005, Figure 2.)

3. Determining the Conversation Type Corresponding to the Natural Language Utterance [1B]/[10B]

The Smith & Hipp-Guinn combination discloses “*determining the conversation type corresponding to the natural language utterance based on whether a user that spoke the natural language utterance has a leader role in an interaction with the voice-enabled device or has a supporter role in the interaction with the voice-enabled device*” [1B] and “*determine the conversation type corresponding to the natural language utterance based on whether a user that spoke the natural language utterance has a leader role in an interaction with the voice-enabled device or has a supporter role in the interaction with the voice-enabled device*” [10B].

As discussed in §III.A, the ’765 patent identifies three conversation types, which, as shown in the table below, differ based on the respective roles of and information allocation between the user and the system as well as the goal to be achieved:

	Participant A User	Participant B Voice User Interface
QUERY		
GOAL	Get information/action	Provide information/action
ROLE	Leader/Consumer	Supporter/Dispenser
INFORMATION ALLOCATION	Less	More
DIDACTIC		
GOAL	Get information	Provide information
ROLE	Follower/Consumer	Leader/Dispenser
INFORMATION ALLOCATION	Less	More
EXPLORATORY		
GOAL	Gather/share information	Gather/share information
ROLE	Follower/Consumer and Leader/Dispenser	Follower/Consumer and Leader/Dispenser
INFORMATION ALLOCATION	Equal or alternating	Equal or alternating

(EX-1001, 15:20-40.)

As discussed above, Smith & Hipp discloses that its voice dialog system can operate in at least four different initiative modes, sometimes called initiative levels. (§IV.A.1; EX-1003, ¶164.) These initiative modes, like the '765 patent's conversation types, are determined on a shifting spectrum based on whether the user has control (*i.e.*, the user is the leader, the system is the supporter) or the system has control (*i.e.*, the system is the leader, the user is the supporter). (EX-1003, ¶164.) As discussed below, Smith & Hipp's four initiative modes correspond closely to the '765 patent's conversation types, except that Smith & Hipp puts a finer point on the spectrum by disclosing what the '765 patent calls an

“exploratory” conversation where the relative role of the user and the system are fluid between system and user:

Control	Smith & Hipp Initiative Mode	'765 Patent Conversation Type
Computer	Directive	Didactic
Both Computer and User	Suggestive	Exploratory, With Bias Towards the System
Both Computer and User	Declarative	Exploratory, With Bias Towards the User
User	Passive	Query

(EX-1003, ¶164.)

Directive Mode = Didactic Conversation: Smith & Hipp’s **directive** mode corresponds to a **didactic** conversation. (EX-1003, ¶165.) According to the ’765 patent, “[i]n a didactic conversation, a leader of the conversation may control information desired by a supporter of the conversation. The supporter’s role may be limited to regulating an overall progression of the conversation and interjecting queries for clarification.” (EX-1001, 15:7-9.) A didactic conversation is led by the computer, with the user in the supporter role. (EX-1001, 15:28-33, claim 5 (“the conversation type ... is determined as a didactic conversation if the user that spoke the natural language utterance has the supporter role”)); EX-1003, ¶165.)

Like a didactic conversation, in Smith & Hipp’s directive mode, the computer controls the conversation as the leader, and the user takes on a supporter

role. (EX-1003, ¶166.) As Smith & Hipp explains, in directive mode, “[t]he **computer has complete dialog control**. It recommends a subgoal for completion and will use whatever dialog is necessary to obtain the needed item of knowledge related to the subgoal.” (EX-1005, 282.) Thus, in directive mode, “[u]nless the user explicitly needs some type of clarification, the computer will select its response solely according to its next goal for the task. If the user expresses need for clarification about the previous goal, this must be addressed first. No interruptions to other subdialogs are allowed.” (EX-1005, 289.)

Suggestive Mode = Exploratory Conversation With Bias Towards the System; Declarative Mode = Exploratory Conversation With Bias Towards the User: According to the ’765 patent, “[i]n an exploratory conversation, both participants share leader and supporter roles, and the conversation may have no specific goal, or the goal may be improvised as the conversation progresses.” (EX-1001, 15:12-16; *see also* EX-1001, claim 5 (“the conversation type ... is determined as an exploratory conversation if the user that spoke the natural language utterance has both the leader role and the supporter role”).)

Smith & Hipp’s **suggestive** and **declarative** modes are functionally equivalent to or render obvious the ’765 patent’s **exploratory** conversation type, as the user and the computer are permitted to each take on a leader or supporter role in those modes. (EX-1003, ¶168.)

In **suggestive** mode, “[t]he computer still has dialog control, but not as strongly. The computer will make suggestions about the subgoal to perform next, but it is also willing to change the direction of the dialog according to stated user preferences.” (EX-1005, 283.) Thus, in suggestive mode, the user and the system may alternate serving the leader and supporter roles, though control weighs in favor of the computer. (EX-1003, ¶169.)

In **declarative** mode, “[t]he user has dialog control, but the computer is free to mention relevant, though not required, facts as a response to the user’s statements.” (EX-1005, 283.) In this mode, the user and the system again may alternate serving the leader and supporter roles, but, in contrast to suggestive mode, control weighs in favor of the user. (EX-1003, ¶170.)

Passive Mode = Query Conversation: The ’765 patent explains that “a query conversation may include a conversational goal of getting a discrete piece of information or performing a particular task, where a leader of the query conversation may have a specific goal in mind and may lead the conversation toward achieving the goal.” (EX-1001, 15:1-5.) The user is the leader in a query conversation, and the computer is the supporter. (EX-1001, 15:23-27, claim 5 (“the conversation type is determined as a query conversation if the user that spoke the natural language utterance has the leader role”); EX-1003, ¶171.)

Similar to the '765 patent's query conversation, in Smith & Hipp's passive mode, the user leads the conversation, *e.g.*, by asking direct questions, and the computer supports the user by providing responses to the user's questions. (EX-1003, ¶172.) Smith & Hipp teaches that, in passive mode, "[t]he user has **complete dialog control**. Consequently, the computer will passively acknowledge user statements. It will provide information only as a direct response to a user question." (EX-1005, 289; *see also* EX-1005, 283 ("The user has complete dialog control. The computer responds directly to user questions and passively acknowledges user statements without recommending a subgoal as the next course of action.").)

Determining the conversation type: The '765 patent claims require that the conversation type be "determined ... based on" a user's role in the conversation at any given time. In other words, as roles and information allocation shift, so too should the conversation type that is being applied by the system. (EX-1003, ¶173.) When the user is the leader and has full control, the system should apply a conversational type that is adaptive to the user's role; here, Smith & Hipp discloses the use of the "passive" initiative mode. (EX-1003, ¶173; EX-1005, 289.) On the other hand, when the user transitions to a supporter role and the system is in control as the leader, the system should operate in Smith & Hipp's "directive" mode. (EX-1003, ¶173; EX-1005, 289.)

Smith & Hipp discusses in theory this shifting conversation types, referred to as “variable initiative dialog.” (EX-1003, ¶174.) Smith & Hipp discloses that a voice dialog system “can participate in variable initiative dialog if it properly manages (1) the selection of the current subdialog, (2) the level of assertiveness in its outputs, and (3) the interpretation of its inputs.” (EX-1005, 289; *see also* EX-1005, 284 (“Our theory also shows how variable initiative is built into the same simple architecture.”).) In fact, Smith & Hipp extensively discusses this variable initiative goal, which “allows either participant to have control, and it allows the initiative to change between participants during the exchange.” (EX-1005, 289.)

While Smith & Hipp discloses the use of various initiative modes corresponding to the conversation types of the ’765 patent, it also acknowledges that its existing system did not include an automatic mechanism for selecting among the variable initiatives. (*See* EX-1005, 316 (noting that “an automatic mechanism is needed for setting the appropriate level of initiative as an interaction proceeds”); *see also* EX-1006, 278 (“Smith presents no algorithm for the computer to change initiative during a dialogue.”).) Guinn provides just such an automatic mechanism for initiative switching for use in the Smith & Hipp system. (EX-1003, ¶176.)

In particular, as discussed in §IV.A.2 above, Guinn discloses an initiative-setting scheme called “Continuous” mode, where “the more knowledgeable agent

(defined by which agent's first-ranked branch is more likely to succeed) is initially given initiative. If that branch fails, this agent's second-ranked branch is compared to the other agent's first-ranked branch with the winner gaining initiative." (EX-1006, 281.)

Applying Guinn's teachings related to Continuous mode to Smith & Hipp's four initiative modes, if the combined Smith & Hipp-Guinn system determines during a conversation that the computer is the more knowledgeable agent, the system gives initiative (or control) to the computer. (EX-1003, ¶177.) In such a case, the system would therefore be operating in **directive** mode, where "[t]he computer has complete dialog control." (EX-1005, 282; EX-1003, ¶178.) On the other hand, if the system determines during a conversation that the user is the more knowledgeable agent, it gives the initiative (control) to the user, thereby operating in **passive** mode, where "[t]he user has complete dialog control." (EX-1005, 283.)

In short, in the combined system of Smith & Hipp-Guinn, the conversation type (initiative mode) is determined based on whether the user ("*a user that spoke cothe natural language utterance*") has more knowledge (and thus the user is the leader of the conversation) or whether the computer has more knowledge (and thus the user is the supporter of the conversation). (EX-1003, ¶¶178-180.)

Accordingly, the Smith & Hipp-Guinn combination discloses limitations [1B]/[10B]. (EX-1003, ¶¶162-81.)

**4. Generating a Response to the Natural Language Utterance
[1C]/[10C]**

The Smith & Hipp-Guinn combination discloses “*generating a response to the natural language utterance with a format based on the conversation type, wherein the format is adapted to limit the user’s future input to interjecting queries or requests for clarification if the user has the supporter role*” [1C] and “*a response builder configured to generate a response to the natural language utterance with a format based on the conversation type, wherein the format is adapted to limit the user’s future input to interjecting queries or requests for clarification if the user has the supporter role*” [10C].

By the priority date of the ’765 patent, spoken dialogue systems had developed strategies to help control the flow of dialog. (EX-1003, ¶¶68-73; §III.B.) One such tool was to generate responses that were formatted to either maintain or relinquish dialog control. (EX-1003, ¶74.) Various systems would use these tools to make it so that control of the dialog flow remained where it was expected: in user-led systems with the user, and in system-led systems with the system. (EX-1003, ¶49.) The ’765 patent is directed to this idea; it claims that when the computer has dialog control (*i.e.*, the user is in the supporter role), that the response is generated in a format that would limit the user’s future input “to interjecting queries or requests for clarification.” (EX-1001, 19:51-52.) In other

words, to limit the user from taking control of the dialog and moving it in another direction. (EX-1003, ¶50.)

The combined system of Smith & Hipp-Guinn discloses this. In that system, where the computer has more knowledge than the user and has initiative (*i.e.* control), the system operates in directive mode where the “computer has complete dialog control.” (EX-1005, 282.) In directive mode, “[u]nless the user explicitly needs some type of clarification, the computer will select its response solely according to its next goal for the task. If the user expresses need for clarification about the previous goal, this must be addressed first. No interruptions to other subdialogs are allowed.” (EX-1005, 289.) In other words, in directive mode, the user’s future input is limited to “*interjecting queries or requests for clarification if the user has the supporter role.*” (EX-1003, ¶183.)

Even more, in contrast to the ’765 patent, Smith & Hipp disclose *how* to format a response in a manner that is adapted to restrict future input. As Smith & Hipp explain, it had been known since at least 1988 that assertive statements (*e.g.*, commands) are most likely to maintain dialog control with the speaker. (EX-1003, ¶184.)

Smith & Hipp provides a sample dialog from its system, confirming that, in directive mode, the user’s input is in fact limited to queries to or requests for clarification from the computer through the use of assertive commands:

1. Computer: Put the knob to one zero.
2. User: Okay.
3. Computer: What is the voltage between connector one two one and connector three four?
4. User: I do not know.
5. Computer: Locate the voltmeter.
6. User: Done.
7. Computer: Add a wire between the minus com hole on the voltmeter and the connector one two one.
8. User: How do I do that?
9. Computer: Connect the end of the black wire with the small plug to the minus corn hole on the voltmeter.
10. User: Done.
11. Computer: Connect the end of the black wire with the large plug to connector one two one.
12. User: Which knob?
13. Computer: Near the top boundary and in the center is a green region. Now, in the middle right of this green region is the control knob.
14. User: Okay.
15. Computer: Put the knob to zero one.
16. User: Done.

17. Computer: Connect the end of the black wire with the large plug to connector one two one.

18. User: Where is connector three four?

19. Computer: Near the left boundary and toward the middle is a green region. Now, in the bottom center of this green region is connector three four.

20. User: Okay.

21. Computer: What is the voltage between connector one two one and connector three four?

22. User: Zero.

(EX-1005, 284-85.)

Accordingly, the Smith & Hipp-Guinn combination discloses limitations [1C]/[10C]. (EX-1003, ¶¶182-185.)

C. Claims 2 and 11 – “format of the response is adapted to provide discrete piece of information, ... a plurality of pieces of information, or obtain information.”

The Smith & Hipp-Guinn combination discloses “[2A] *determining the conversation type comprises determining a conversational goal corresponding to the natural language utterance, wherein [2B] the format of the response is adapted to provide a discrete piece of information to the user, to provide a plurality of pieces of related information to the user, or to obtain one or more pieces of information from the user*” [2]. (EX-1003, ¶¶186-205.) The limitations of claims 2

and 11 are identical in substance, so Petitioner addresses them together here. (EX-1003, ¶186.)

1. Determining a Conversational Goal Corresponding to the Natural Language Utterance [2A]/[11A]

The Smith & Hipp-Guinn combination discloses “*determining a conversational goal corresponding to the natural language utterance.*” (EX-1003, ¶¶187-93.) Smith & Hipp teaches that an effective dialog system must interpret each user utterance in light of the user’s goal or purpose. (EX-1003, ¶187.) Smith & Hipp expressly describes its system as “goal-oriented” (EX-1005, 285), where efficient dialog requires that each participant (including the computer) understand “the purpose of the interaction” (EX-1005, 282). In other words, the Smith & Hipp system must determine the conversational goal underlying the user’s utterance in order to decide how to respond. (EX-1003, ¶187.)

Smith & Hipp further discloses that the dialog controller—*i.e.*, “the overall ‘boss’ of the dialog processing system”—formulates goals at the top level to be passed on to the theorem-proving stage.” (EX-1005, 297.) Smith & Hipp explains the following basic cycle performed by the dialog controller:

- (1) Obtain suggested goal from the domain processor.
- (2) Based on the suggested goal and the current state of the dialog, select the next goal to be pursued by the computer and determine the expectations associated with that goal. (The goal may thus be selected

from one of the active subdialogs. The choice is partially dependent on the current level of initiative.)

(3) Attempt to complete the goal using the IPSIM system, possibly invoking voice interactions.

(4) Update system knowledge based on efforts at goal completion.

(5) Determine next operations to be performed by the domain processor in providing a suggested goal.

(6) Go to step 1.

(EX-1005, 297.)

Guinn, like Smith & Hipp, also discloses determining and using the conversational goal. (EX-1003, ¶¶191-93.) For example, Guinn explains that “initiative levels for each goal are defined during the interaction based on 1) explicit and implicit initiative-changing utterances and 2) competency evaluation,” and that “an agent may have initiative over one goal but not another. As goals get pushed and popped from the problem-solving stack, initiative changes accordingly.” (EX-1006, 279.) Thus, according to Guinn, “many initiative changes are done implicitly based on which goal is being solved.” (EX-1006, 279.) A POSITA would understand from Guinn that, in order to determine initiative for a conversation (*i.e.*, the conversation type), the system must first understand the conversational goal. (EX-1003, ¶191.)

Accordingly, a POSITA would understand the Smith & Hipp-Guinn combination discloses “*determining a conversational goal corresponding to the natural language utterance.*” (EX-1003, ¶¶187-93.)

2. Format of the Response Is Adapted to Provide ... [2B]/[11B]

The Smith & Hipp-Guinn combination also discloses adapting the format of the response “*to provide a discrete piece of information to the user, to provide a plurality of pieces of related information to the user, or to obtain one or more pieces of information from the user.*” (EX-1003, ¶¶194-205.)

Smith & Hipp discloses adjusting the format and content of the system’s response depending on the user’s needs and the dialog context. (EX-1003, ¶194.)

For example, Smith & Hipp discloses:

The system must output statements that are compatible with its level of initiative. This means that the output generator must have a parameter that enables the system to specify assertiveness. Examples of the request to ‘turn the switch up’ at various levels of assertiveness are as follows:

Turn the switch up.	(command)
Would you [please] turn the switch up?	(request)
Can you turn the switch up?	(indirect request)
The switch can be turned up.	(statement of fact)
Turning the switch up is necessary.	(statement of fact)

Two examples of querying for the switch position are as follows:

What is the switch position? (request)

I need to know the switch position. (indirect request)

(EX-1005, 289-90.) Thus, Smith & Hipp teaches a response format adaptation based on four levels of assertiveness as command, request, indirect request, and statement of fact. (EX-1003, ¶195.) A POSITA would understand that Smith & Hipp's system generates responses that can provide discrete pieces of information in the "statement of fact" assertive level, or obtain information from the user through "request" or "indirect request" assertive levels.

Response format adapted to provide a discrete piece of information: As noted above, Smith & Hipp's system may operate in passive mode, where the computer "will provide information only as a direct response to a user question." (EX-1005, 289; *see also* EX-1005, 283 ("The computer responds directly to user questions").) In other words, in passive mode, the computer provides only the specific information requested by the user, without volunteering additional details. (EX-1003, ¶196.) Under such circumstances, the system's response is adapted to "*provide a discrete piece of information to the user.*" (EX-1003, ¶196.)

Smith & Hipp's system provides a discrete piece of information to the user in other initiative modes as well.¹ For example, Smith & Hipp provides the following sample dialog of the system operating in declarative mode:

1. Computer: This is the circuit-fixit-shop. How may I help you?
2. User: RS111 is not working.
3. Computer: I am familiar with that circuit. The LED is supposed to be displaying alternately flashing one and seven.
4. User: There is no wire on one zero four.
5. Computer: There is supposed to be a wire between connector one zero four and connector one zero two.
6. User: The circuit is working.
7. Computer: Glad to have been of assistance. Good-bye.

(EX-1005, 285.) In the underlined sentence, the system's response provides a discrete piece of information responsive to the user and relevant to the conversational goal. (EX-1003, ¶197.)

¹ Depending on the context of the conversation, the system may also provide short and direct responses in other initiative modes as well, particularly where the user is in control of the dialog (*e.g.*, in declarative mode). In other words, the Smith & Hipp system does not only “*provide a discrete piece of information to the user*” in passive mode. (EX-1003, ¶197.)

Response format adapted to provide a plurality of pieces of related information: As discussed above, Smith & Hipp also discloses a declarative mode, in which the computer is “free to mention relevant, though not required, facts as a response to the user’s statements.” (EX-1005, 289.) In other words, in declarative mode, the computer goes beyond a simple answer and include additional related facts in its response to the user. (EX-1003, ¶198.) Similarly, in directive mode, the computer is in complete control and “will use whatever dialog is necessary to obtain the needed item of knowledge related to the subgoal.” (EX-1005, 282.) Thus, in directive mode, the computer can provide multiple pieces of information to the user for the user to assist in achieving the goal. (EX-1003, ¶199.) Each of these types of responses is “*adapted ... to provide a plurality of pieces of related information to the user.*” (EX-1003, ¶¶199-200.)

For example, the following is an excerpt from a sample dialog where the system is operating in directive mode:

11. Computer: Connect the end of the black wire with the large plug to connector one two one.
12. User: Which knob?
13. Computer: Near the top boundary and in the center is a green region. Now, in the middle right of this green region is the control knob.
14. User: Okay.
15. Computer: Put the knob to zero one.

16. User: Done.

17. Computer: Connect the end of the black wire with the large plug to connector one two one.

18. User: Where is connector three four?

19. Computer: Near the left boundary and toward the middle is a green region. Now, in the bottom center of this green region is connector three four.

(EX-1005, 284-85.) In each of the underlined responses, the system provides multiple pieces of related information in response to the user. (EX-1003, ¶201.)

Response format adapted to obtain information: The Smith & Hipp system also can provide a response “*to obtain one or more pieces of information from the user.*” (EX-1003, ¶¶201-203.) For example, in directive mode, the computer “recommends a subgoal for completion and will use whatever dialog is necessary **to obtain the needed item of knowledge** related to the subgoal.” (EX-1005, 282.) Smith & Hipp provides sample responses from the system designed to obtain information, such as: “What is the voltage between connector one two one and connector three four?” or “Where is connector three four?” (EX-1005, 284-85.)

Accordingly, the Smith & Hipp-Guinn combination renders obvious claims 2 and 11. (EX-1003, ¶¶186-205.)

D. Claim 3 – “discrete” responses shorter than “plurality of pieces” responses

The Smith & Hipp-Guinn combination discloses “*the response generated in the format that is adapted to provide the discrete piece of information is shorter than the response generated in the format that is adapted to provide the plurality of pieces of related information*” [3]. (EX-1003, ¶206.) In short, a POSITA would understand this limitation to mean that a brief answer providing a single fact should be shorter in length than a more detailed answer providing multiple facts. (EX-1003, ¶207.) That a response with one fact can be shorter than a response with multiple facts is a natural and obvious result, as confirmed by Smith & Hipp’s disclosures. (EX-1003, ¶207.)

As discussed above, Smith & Hipp discloses that, in passive mode, “[the] computer will ... provide information **only as a direct response to a user question.**” (EX-1005, 289.) In other words, the computer provides only the single piece of information requested by the user and nothing more. (EX-1003, ¶¶207-208.) A POSITA would understand that such a response is typically a single sentence or a brief statement confined to the specific answer. (EX-1003, ¶208.) On the other hand, where the system is less limited in how it can respond (*e.g.*, in declarative mode, where “[t]he user has dialog control, but the computer is free to mention relevant, though not required, facts as a response to the user’s statements”

(EX-1005, 283), it is able to provide additional related facts, and the response naturally becomes longer. (EX-1003, ¶208.)

As noted above, Smith & Hipp provides examples where the system responds with a discrete piece of information. For example, the response, “There is supposed to be a wire between connector one zero four and connector one zero two.” provides one discrete piece of information and consists of 17 words. (EX-1005, 285.)

Smith & Hipp also provides examples where the system responds with a plurality of pieces of related information. For example, the response “Near the top boundary and in the center is a green region. Now, in the middle right of this green region is the control knob.” provides multiple pieces of related information and naturally consists of more words (25). (EX-1005, 285.) This response is longer than the response providing a single piece of information. (EX-1003, ¶¶209-210.)

Accordingly, the Smith & Hipp-Guinn combination discloses “*the response generated in the format that is adapted to provide the discrete piece of information is shorter than the response generated in the format that is adapted to provide the plurality of pieces of related information.*” (EX-1003, ¶¶206-211.)

E. Claims 4 and 12 – “determining conversation type” based on “conversational goal.”

The Smith & Hipp-Guinn combination discloses “[4A] *determining the conversation type includes determining a conversational goal corresponding to the*

natural language utterance, wherein [4B] the conversation type is determined as a query conversation if the conversational goal is to gather discrete information or perform a discrete task, is determined as a didactic conversation if the conversational goal is to gather related information, or is determined as an exploratory conversation if the natural language utterance has no specific goal.”

(EX-1003, ¶¶212-21.) The limitations of claims 4 and 12 are identical in substance, so Petitioner addresses them together here. (EX-1003, ¶212.)

1. Determining a Conversational Goal Corresponding to the Natural Language Utterance [4A]/[12A]

The Smith & Hipp-Guinn combination discloses “*determining a conversational goal corresponding to the natural language utterance.*” (EX-1003, ¶¶214-16.) This limitation is satisfied for the same reasons as for identical limitation [2A] above. (See §IV.C.1.)

2. Conversation Type Is Determined to Be a Query, Didactic, or Exploratory Conversation [4B]/[12B]

“*Query*” Conversation: The Smith & Hipp-Guinn combination discloses that “*the conversation type is determined as a query conversation if the conversational goal is to gather discrete information or perform a discrete task.*” (EX-1003, ¶218.) As noted above, Smith & Hipp’s passive mode corresponds to the query conversation type of the ’765 patent. (§IV.B.3.) Smith & Hipp discloses that, in passive mode, “[t]he computer responds directly to user questions ...

without recommending a subgoal as the next course of action.” (EX-1005, 283.) In light of this disclosure, a POSITA would understand that the conversational goal in passive mode, as for a query conversation, is “*to gather discrete information or perform a discrete task.*” (EX-1003, ¶218.)

“*Didactic*” Conversation: The Smith & Hipp-Guinn combination discloses that “*the conversation type ... is determined as a didactic conversation if the conversational goal is to gather related information.*” (EX-1003, ¶219.) As noted above, Smith & Hipp’s directive mode corresponds to the didactic conversation type of the ’765 patent. (§IV.B.3.) Smith & Hipp discloses that, in directive mode, the computer “will use whatever dialog is necessary to obtain the needed item of knowledge related to the subgoal.” (EX-1005, 282.) The conversational goal in directive mode is therefore “*to gather related information.*” (EX-1003, ¶219.)

“*Exploratory*” Conversation: The Smith & Hipp-Guinn combination discloses that “*the conversation type ... is determined as an exploratory conversation if the natural language utterance has no specific goal.*” (EX-1003, ¶220.) As noted above, Smith & Hipp’s suggestive and declarative modes correspond to the exploratory conversation of the ’765 patent, with the difference between the two modes being whether the control bias is in the user’s favor (declarative) or the computer’s favor (suggestive). (See §IV.B.3; EX-1003, ¶220.) In both modes, there is no specific goal, just like an exploratory conversation. (EX-

1003, ¶220.) For example, in suggestive mode, “the computer will make suggestions about the subgoal to perform next, but it is also willing to change the direction of the dialog according to stated user preferences.” (EX-1005, 283.)

F. Claims 5 and 13 – “conversation type” determined based on “role”

The Smith & Hipp-Guinn combination discloses “*the conversation type is determined as a query conversation if the user that spoke the natural language utterance has the leader role, is determined as a didactic conversation if the user that spoke the natural language utterance has the supporter role, or is determined as an exploratory conversation if the user that spoke the natural language utterance has both the leader role and the supporter role.*” The limitations of claims 5 and 13 are identical in substance, so Petitioner addresses them together here. (EX-1003, ¶222.)

Claims 5 and 13 provide a simple mapping between dialog control (*i.e.*, whether the user or system controls or “leads” the dialog) and the type of conversation. That mapping is as follows:

“If the user that spoke the natural language utterance has:	Conversation Type is Then:
“the leader role”	“Query”
“the supporter role”	“Didactic”
“both the leader role and the supporter role”	“Exploratory”

As will be explained in more detail in the sections that follow, this mapping corresponds directly with the mapping provided by Smith & Hipp's initiative modes.

'765 Role	'765 Conversation Type	S&H Dialog Control	S&H Initiative Mode
"the leader role"	"Query"	User Control	Passive
"the supporter role"	"Didactic"	System Control	Directive
"both the leader role and the supporter role"	"Exploratory"	Alternating Control	Suggestive/Declarative

Starting with the so-called "Query" mode, as discussed above, the '765 patent explains that "a query conversation may include a conversational goal of getting a discrete piece of information or performing a particular task, where a leader of the query conversation may have a specific goal in mind and may lead the conversation toward achieving the goal." (EX-1001, 15:1-5.) The user is the leader in a query conversation, and the computer is the supporter. (EX-1001, 15:23-27)

The Smith & Hipp-Guinn combination renders the link between the user leading the conversation and the "Query" mode obvious. Smith & Hipp teaches that when the user controls the dialogue (*i.e.* the user is taking the lead), the system operates in a completely passive mode, essentially following the user's queries and instructions. (EX-1005, 282-83.) As Smith & Hipp describes, in the passive mode,

“[t]he user has complete dialog control. The computer responds directly to user questions and passively acknowledges user statements without recommending a subgoal as the next course of action.” (EX-1005, 283.) A POSITA would immediately understand that in Smith & Hipp’s passive mode, the user is the “leader” of the conversation, as that term is described in the ’765 patent. (EX-1003, ¶224.)

Same for the so-called “Didactic” mode. As discussed above, the ’765 patent explains that “[i]n a didactic conversation, a leader of the conversation may control information desired by a supporter of the conversation. The supporter’s role may be limited to regulating an overall progression of the conversation and interjecting queries for clarification.” (EX-1001, 15:7-9.) A didactic conversation is led by the computer, with the user in the supporter role. (EX-1001, 15:28-33, claim 5 (“the conversation type ... is determined as a didactic conversation if the user that spoke the natural language utterance has the supporter role”); EX-1003, ¶225.)

The Smith & Hipp-Guinn combination renders the link between the computer leading the conversation (*i.e.*, the user “supporting”) and the “didactic” mode obvious. In its highest system-initiative mode (described as “directive”), “the computer has complete dialog control” and directs the course of the conversation to achieve a subgoal. (EX-1005, 282.) The user in that mode is essentially a supporter, following the system’s prompts or instructions. (EX-1003, ¶225.) The

nature of this interaction is thus didactic: the system is imparting information or instructing the user to accomplish tasks, much like a teacher or guide. (EX-1003, ¶225.) A POSITA would equate Smith & Hipp’s system-led dialog mode with a didactic conversation, since the goal in such mode is to convey information to the less-informed party (the user) in a controlled manner. Therefore, when the user occupies the supporter role (and the system leads), Smith & Hipp makes it obvious that the conversation is of a directive (or in the lingo of the ’765 patent, didactic) type. (EX-1005, 282.)

Finally, Smith & Hipp also renders obvious an “exploratory” conversation where the user alternates between both the leader role and the supporter role. According to the ’765 patent, “[i]n an exploratory conversation, both participants share leader and supporter roles, and the conversation may have no specific goal, or the goal may be improvised as the conversation progresses.” (EX-1001, 15:12-16; *see also* EX-1001, claim 5 (“the conversation type ... is determined as an exploratory conversation if the user that spoke the natural language utterance has both the leader role and the supporter role”).)

Smith & Hipp’s model of variable initiative dialogue makes clear that the user and system may alternate control depending on the user’s knowledge and behavior. Smith & Hipp describes distinct modes—suggestive and declarative—in which initiative is permitted to shift dynamically. (EX-1003, ¶226.) In suggestive

mode, “[t]he computer still has dialog control, but not as strongly. The computer will make suggestions about the subgoal to perform next, but it is also willing to change the direction of the dialog according to stated user preferences.” (EX-1005, 283.) Thus, in suggestive mode, the user and the system may alternate serving the leader and supporter roles, though control weighs in favor of the computer. (EX-1003, ¶226.) In declarative mode, “[t]he user has dialog control, but the computer is free to mention relevant, though not required, facts as a response to the user’s statements.” (EX-1005, 283.) In this mode, the user and the system again may alternate serving the leader and supporter roles, but, in contrast to suggestive mode, control weighs in favor of the user. (EX-1003, ¶227.)

Disclosures in Guinn reinforces the concept of dynamically alternating dialog control between users and the system, stating that “dialogue initiative should always pass immediately to the participant who is best able to handle the current task.” (EX-1006, 278; EX-1003, ¶228.) This dynamic allows users to lead some goals and support others—even within the same sequence of utterances. (EX-1003, ¶228.) Guinn formalizes this through explicit and implicit initiative-changing utterances, explaining that users can seize control by posing new questions or revert to a supporting role when the system’s knowledge is superior. (EX-1006, 279.)

Guinn’s model specifically highlights the role of questions and answers in shifting control: when one agent asks another to satisfy a goal, the second agent gains initiative over that goal and its subgoals until control is passed back. (EX-1006, 279.) A POSITA would readily understand from Guinn, as well as Smith & Hipp, that an alternating initiative model inherently results in a user who sometimes leads and sometimes supports within a single dialog flow.

Accordingly, the Smith & Hipp-Guinn combination renders obvious claims 5 and 13. (EX-1003, ¶¶222-230.)

G. Claims 6, 7, and 14 – “conversation type” determined based on “allocation of information.”

The Smith & Hipp-Guinn combination discloses “*determining the conversation type [is] based on an allocation of information between the user and the speech engine*” [6][14] and “*determining the conversation type further comprises determining an allocation of information between the user and the speech engine*” [7]. In addition, the Smith & Hipp-Guinn combination further discloses “*the conversation type is determined to be a query conversation or a didactic conversation if the user that spoke the natural language utterance has a lower allocation of information related to the natural language utterance than the speech engine, or is determined as a didactic conversation if the user and the speech engine have an equal or alternating allocation of information [related to the natural language utterance].*” [6][14] The overlapping limitations of claims 6,

7, and 14 are identical in substance, so Petitioner addresses them together here.

(EX-1003, ¶231.)

Starting with “determining the conversation type based on an allocation of information between the user and the speech engine” (claims 6, 7, and 14), Smith & Hipp discloses that conversation type is directly related to the allocation of information between the user and the speech engine. (EX-1003, ¶234.) For example, in discussing the ideal conversation type to create an “efficient interaction,” Smith & Hipp notes that a spoken dialog system should “yield control so that the more competent partner can lead the way to the fastest possible solution.” (EX-1005, 282.) Conversely, if the user lacks information, the system “must be able to carry out its own problem-solving process and direct the actions to task completion.” (EX-1005, 282.) Smith & Hipp does not claim to have invented this concept; indeed it is a foundational idea in spoken-dialog systems. (EX-1003, ¶¶238-39.)

Because of the importance of information allocation in dialog control, Smith & Hipp discloses in detail various tools for tracking the information allocation dynamically during the course of a conversation and maintaining this representation in the “user model.” (EX-1003, ¶235.) Moreover, Smith & Hipp discloses how updates to the user model should affect the conversation type,

describing how there should be a “change of initiative” based on which party holds the requisite knowledge to advance the task at hand. (EX-1003, ¶237.)

While the Smith & Hipp system, like the ’765 patent, did not disclose an automatic process by which conversation type was updated mid-conversation to reflect an allocation of information, Guinn provides that disclosure if it is needed. Namely, as discussed above, Guinn discloses multiple suitable algorithms for assigning initiative modes (*i.e.*, conversation types) based on which participant is the “more knowledgeable agent” as defined by, for example, “which agent has the greater total percentage of knowledge.” (EX-1003, ¶241.) Thus the Smith & Hipp-Guinn combination fully discloses the limitation in claims 6, 7, and 14 of “determining the conversation type based on an allocation of information between the user and the speech engine.”

The Smith & Hipp-Guinn combination also fully discloses the specific mapping between information allocation and conversation type disclosed in the ’765 patent. That mapping, as written in the claims, is depicted below.

Allocation of Information	Conversation Type
IF “user that spoke the natural language utterance has a lower allocation of information related to the natural language utterance than the speech engine”	THEN “Query or Didactic”
IF “user and the speech engine have an equal or alternating allocation of information related to the natural language utterance”	THEN “Didactic”

As an initial matter, a POSITA reading the '765 patent would conclude that the mapping in the claim language between “didactic” and cases where the “user and the speech engine have an equal or alternating allocation of information related to the natural language utterance” to either be a drafting error or to make no sense. (EX-1003, ¶244.) As written, the claim provides two different, and arguably contradictory definitions of didactic, and moreover defines didactic using the exact same language used to define “exploratory” in the specification. (EX-1003, ¶¶232-233.) After accounting for this drafting error, the mapping between the allocation of information and the conversation type is as follows:

Allocation of Information	Conversation Type
IF “user that spoke the natural language utterance has a lower allocation of information related to the natural language utterance than the speech engine”	THEN “Query or Didactic”
IF “user and the speech engine have an equal or alternating allocation of information related to the natural language utterance”	THEN “ <u>Exploratory</u> ”

The Board ultimately need not address the claim element with the drafting error if it is not inclined to do so since, because that element of the claim is written in the conditional, the Board need only conclude that the first condition is obvious over the Smith & Hipp-Guinn combination to find the claim invalid. *See Cybersettle, Inc. v. Nat’l Arbitration Forum, Inc.*, 243 F.App’x 603 (Fed. Cir. 2007) (“[i]f the condition for performing a contingent step is not satisfied, the

performance recited by the step need not be carried out in order for the claimed method to be performed.”). In all events both conditions are rendered obvious over the Smith & Hipp-Guinn combination and are addressed, in turn, below.

First, the Smith & Hipp-Guinn combination discloses the condition that if the user has a lower allocation of information related to the utterance than the speech engine, the conversation type is determined to be a query conversation or a didactic conversation. Smith & Hipp plainly teaches this scenario. (EX-1003, ¶244.) As Smith & Hipp discloses, if the user is relatively uninformed (lacking certain knowledge) and the system holds that information, two outcomes are possible: (1) the user can ask the system for the information, or (2) the system can tell/teach the user the information. Smith & Hipp accommodates both cases.

Specifically, Smith & Hipp notes that the system can either provide information merely in response to user questions (query / passive scenario), or that the system can take control (*i.e.*, initiative) so that it may suggest or explain information needed to reach the goal (didactic/directive scenario). (EX-1005, 282-83.; EX-1003, ¶109-111.) A POSITA would therefore find it obvious that when the user has less information than the system on a subject, the dialogue naturally falls into either a query pattern (user questioning a knowledgeable system) or a didactic pattern (system instructing an uninformed user). Smith & Hipp supplies the reasoning and examples to arrive at this conclusion. (EX-1003, ¶244.)

Second, the Smith & Hipp-Guinn combination renders obvious that if there is an equal or alternating allocation of information then the conversation type should be designated as “exploratory.” (EX-1003, ¶246.)

The specification defines exploratory conversation types as those where “both participants share leader and supporter roles, and the conversation may have no specific goal, or the goal may be improvised as the conversation progresses.” (EX-1001 at 15:13-16.) Put differently, an “exploratory” conversation conceives of dialog control alternating between the user and the system fluidly and with the conversational goal changing as the conversation progresses.

This is exactly what Smith & Hipp discusses in the context of “variable initiative dialogs”, which it notes had been studied for many years even before Smith & Hipp was published. (EX-1005, 282; EX-1003, ¶190.) As Smith & Hipp describes, it is a feature of everyday “efficient” conversation that often “[a] subdialog is opened, leading to another, then another, then a jump to a previously opened subdialog, and so forth, in an unpredictable order until the necessary subgoals have been solved for an overall success.” (EX-1005, 282.). As Smith & Hipp further explains, as subdialogs and goals change during the course of a conversation, so does the allocation of information between the user and the machine as it relates to the given subgoal or on the global task. (EX-1005, 282.). And, as Smith & Hipp also explains, when the allocation shifts, “efficient

interaction requires” that control shifts as well. (EX-1005, 282.). Smith & Hipp thus renders obvious this claimed linking between an equal or alternating allocation of information and the claimed “exploratory” conversation style. (EX-1003, ¶¶246-51.)

What is more, Smith & Hipp’s “suggestive” and “declarative” modes are expressly built to facilitate this type of dynamic variation in relative information allotment and dialog control. (EX-1003, ¶249.) As Smith & Hipp explains, these “intermediate levels of control” allow one participant to “gently rather than strongly lead the interactions.” This facilitates changes in initiative, or control, as in so-called “exploratory” conversation types. (EX-1003, ¶249.)

Again, Guinn provides additional disclosures that demonstrate the connection between the condition of equal or alternating allocations of knowledge the need for dynamic dialog control, providing multiple algorithms that link, in real time, relative allocation of knowledge to the role of the leader of the conversation. (EX-1003, ¶¶248, 251.) Guinn emphasizes the dynamic and fluid nature of these control shifts, noting “efficient human-computer dialogue requires immediate utterance-by-utterance accommodation to the needs of the interaction. ... This means that the dialogue initiative should always pass immediately to the participant who is best able to handle the current task.” (EX-1006, 278.).

Accordingly, the Smith & Hipp-Guinn combination renders obvious claims 6, 7, and 14. (EX-1003, ¶¶231-252.)

H. Claim 8 - “only provides information requested ... limits the user’s future input”

The Smith & Hipp-Guinn combination discloses “*if the conversation type is a query conversation, a response that only provides information requested by the natural language utterance, and generating, if the conversation type is a didactic conversation, a response that controls the conversation and limits the user’s future input to a request or query for clarification*” [8]. (EX-1003, ¶252.)

Breaking this claim into its two parts, it provides that (1) in query (*i.e.*, user-led) conversations, the system only provides in response to the user what was asked for, and (2) in didactic (*i.e.*, system-led) conversations, the system responses should be configured to maintain control of the conversation (*i.e.*, “limit the user’s future input to a request or query for clarification.”) Both of these elements to the claim are directed to well-understood and routine concepts in dialog management. They are also disclosed in Smith & Hipp, as set forth in turn below.

If Query only provide what was asked. As discussed previously, Smith & Hipp’s “passive” mode corresponds to the ’765 patent’s “query” mode; in both the user maintains total control of the dialog. (§IV.B.3; EX-1003, ¶254.) Consistent with user-led spoken dialog systems in the art, (EX-1003, ¶49,) Smith & Hipp discloses that when the user is in control, the job of the system is to “provide

information only as a direct response to a user question.” (EX-1003, ¶254) (EX-1005, 289.)

If Didactic format response to maintain control. On the flip side, as also discussed previously, Smith & Hipp’s “directive” mode corresponds to a system-led scenario. (§IV.A.1; EX-1003, ¶225.) In directive mode, “[t]he computer has complete dialog control” and will drive the conversation by pursuing its own task goals, “select[ing] its response solely according to its next goal for the task,” unless the user needs clarification. (EX-1005, 289.) The system does not allow the user to deviate the dialog except “[i]f the user expresses need for clarification.” (EX-1005, 289.) The reason for this limitation is explained in the background literature, which explains that due to technical limitations in spoken dialog systems, it was critical for tutorial systems to maintain dialogue control due to system limitations. (EX-1003, ¶¶56, 74.)

Accordingly, the Smith & Hipp-Guinn combination renders obvious claim 8. (EX-1003, ¶¶252-257.)

I. Claim 9 - “user controls the conversation or ... provides information”

The Smith & Hipp-Guinn combination discloses “*the natural language utterance is part of a conversation that comprises related natural language utterances from the user*” and “*determining whether the user has the leader role or the supporter role comprises determining whether the user controls the*

conversation or whether the user provides information to the speech engine during the conversation” [9]. (EX-1003, ¶258.)

Claim 9 comprises two subparts: (1) a requirement for what is known as a “multi-turn conversation” (*i.e.*, “conversation that comprises related natural language utterances”) and (2) a requirement for deciding the user’s role based on whether the user *controls* the dialog or whether the user *provides information* during the dialog. The Smith & Hipp-Guinn combination discloses both.

Starting with the requirement for a multi-turn conversation comprising related natural language utterances, Smith & Hipp, for instance, explicitly discloses how “efficient human–machine dialog” is “segmented into utterance sequences, *subdialogs*, that are individually aimed at achieving relevant subgoals.” (EX-1003, ¶259.) Moreover, Smith & Hipp discloses numerous “sample dialogs” comprising conversations between the user and the system where the utterances from the user are “related” insofar as they are all directed towards the target goal. (EX-1003, ¶260.) For example, in the sample dialog below the utterances from the user are all related to the target goal of setting a particular knob to 10 and then a voltage measurement should be made.

1. Computer: Put the knob to one zero.
2. User: Okay.

3. Computer: What is the voltage between connector one two one and connector three four?
4. User: I do not know.
5. Computer: Locate the voltmeter.
6. User: Done.
7. Computer: Add a wire between the minus com hole on the voltmeter and the connector one two one.
8. User: How do I do that?
9. Computer: Connect the end of the black wire with the small plug to the minus corn hole on the voltmeter.
10. User: Done.
11. Computer: Connect the end of the black wire with the large plug to connector one two one.
12. User: Which knob?
13. Computer: Near the top boundary and in the center is a green region. Now, in the middle right of this green region is the control knob.
14. User: Okay.
15. Computer: Put the knob to zero one.
16. User: Done.
17. Computer: Connect the end of the black wire with the large plug to connector one two one.
18. User: Where is connector three four?

19. Computer: Near the left boundary and toward the middle is a green region. Now, in the bottom center of this green region is connector three four.

20. User: Okay.

21. Computer: What is the voltage between connector one two one and connector three four?

22. User: Zero.

(EX-1005, 284-85; EX-1003, ¶184.)

The Smith & Hipp-Guinn combination also discloses deciding the user's "role" based on whether the user controls the dialog or whether the user provides information during the dialog. To recap: the claims and specification of the '765 patent envision two primary roles for a user: "leader" and "supporter." The Smith & Hipp-Guinn combination plainly discloses that a determination of who is "leading" a conversation is directly a function of which participant is controlling the conversation or which participant is providing information in a conversation.

Starting with "control," as discussed above, Smith & Hipp disclose four "levels of initiative" based on which participant "lead[s] the interaction[]." (§V.A.1; EX-1003, ¶¶108-111.) At one extreme, in the directive mode, "the computer has complete dialog control." (EX-1003, ¶109.) At the other, in the passive mode, "the user has complete dialog control." (EX-1003, ¶110.) In the middle, the suggestive mode provides that the computer will "gently rather than

strongly lead the interactions,” whereas for the declarative mode it is the opposite. (EX-1003, ¶111.) These modes illustrate that the user’s role—as either a leader or a supporter—can be inferred from whether the user is initiating dialog acts and driving the interaction forward (leader), or merely responding to system prompts (supporter). (EX-1003, ¶261.)

Similarly, Smith & Hipp discloses that the user’s role (*i.e.*, leader or supporter/in control or not) can be “determined” based on the level of control. Specifically, Smith & Hipp describes a “change of initiative” where the machine should “yield to the user’s control and respond ... to his or her requests.” (EX-1005, 282.) A POSITA would understand this disclosure to mean that if a user is asserting “control” it should update the initiative mode to reflect the “role” of the user. (EX-1003, ¶262.)

Turning next to determining role based on whether the user “provides information” to the system during the conversation, the Smith & Hipp-Guinn combination discloses this too. As discussed above, Smith & Hipp’s four initiative modes reflect the correspondence between the participant controlling the conversation (*i.e.*, leading or supporting) and whether the user is providing or receiving information from the system. (§IV.B.3; EX-1003, ¶263.).

Likewise, Smith & Hipp discloses that the user’s role (*i.e.*, leader or supporter/in control or not) can be “determined” based on whether the user is

providing information to the system. Specifically, Smith & Hipp explains that if the user is more capable or has more relevant knowledge for a given task, the machine should “yield control so that the more competent partner can lead” the dialogue to a solution. (EX-1003, ¶263.) Guinn likewise explains questions and answers are “two specific dialogue cues” for determining who is leading a conversation. (EX-1003, ¶265.)

Accordingly, the Smith & Hipp-Guinn combination renders obvious claim 9. (EX-1003, ¶¶258-66.)

V. Discretionary Denial is Not Appropriate

Pursuant to the Stewart Memorandum dated March 26, 2025, titled “Interim Process for PTAB Workload Management,” Petitioner will respond to any discretionary denial arguments PO may raise through the Board’s bifurcated briefing process.

Nevertheless, to simplify the *Fintiv* analysis, and following the Board’s precedential decision in *Sotera Wireless, Inc. v. Masimo Corp.*, IPR2020-01019, Paper 12 at 18-19 (PTAB Dec. 1, 2020), Petitioner stipulates that, if IPR is instituted, Petitioner will not pursue in the related district court proceeding any ground that Petitioner raised or reasonably could have raised against the challenged claims during the instituted IPR.

VI. Mandatory Notices

A. Real Party In Interest

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

B. Related Matters

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

- *VB Assets, LLC v. Samsung Electronics Co. Ltd. et al.*, 2:24-cv-00828-JRG-RSP (E.D. Tex. filed Oct. 9, 2024)
- *VB Assets, LLC v. Apple, Inc.*, No. 1:24-cv-01368-MN (D. Del., filed Dec. 13, 2024)
- *Samsung Electronics Co. Ltd. et al. v. VB Assets, LLC*, IPR2025-00866 (PTAB)

C. Notice of Counsel and Service Information

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

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

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

VII. Conclusion

IPR is respectfully requested.

Date: April 25, 2025

Respectfully submitted,

/Lori A. Gordon/

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

Lead Counsel for Petitioner

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

Pursuant to 37 C.F.R. §42.24(a), Samsung hereby certifies that portions of the above-captioned Petition for *Inter Partes* Review of U.S. Patent No. 8,515,765 in accordance with and reliance on the word count provided by the word-processing system used to prepare this Petition, that the number of words in this paper is 12,418. Pursuant to 37 C.F.R. §42.24(a), this word count is in compliance and excludes the table of contents, table of authorities, mandatory notices under §42.8, certificate of service, certificate of word count, appendix of exhibits, and any claim listing. This word count was prepared using Microsoft Word.

Date: April 25, 2025

Respectfully submitted,

/Lori A. Gordon/

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

Lead Counsel for Petitioner

CERTIFICATE OF SERVICE

The undersigned certifies that a true copy of the Petition for *Inter Partes* Review of U.S. Patent No. 8,515,765 together with all exhibits identified in the above Table of Exhibits and Petitioners' Power of Attorney, have been served on the Patent Owner via FedEx Next Business Day Delivery on the below date, at the correspondence address of record as listed on the Patent Center:

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

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

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

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

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

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

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

VBA-Samsung@alston.com

Attorneys for VB Assets LLC

Date: April 25, 2025

/Lori A. Gordon/

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

Lead Counsel for Petitioner

APPENDIX A – LIST OF CLAIMS

[1P] A method of providing a voice interface, comprising:

[1A] receiving, at a speech engine, a natural language utterance from a voice-enabled device, the natural language utterance corresponding to a conversation type;

[1B] determining the conversation type corresponding to the natural language utterance based on whether a user that spoke the natural language utterance has a leader role in an interaction with the voice-enabled device or has a supporter role in the interaction with the voice-enabled device; and

[1C] generating a response to the natural language utterance with a format based on the conversation type, wherein the format is adapted to limit the user's future input to interjecting queries or requests for clarification if the user has the supporter role.

[2A] The method recited in claim 1, wherein determining the conversation type comprises determining a conversational goal corresponding to the natural language utterance,

[2B] wherein the format of the response is adapted to provide a discrete piece of information to the user, to provide a plurality of pieces of related information to the user, or to obtain one or more pieces of information from the user.

[3] The method recited in claim 2, wherein the response generated in the format that is adapted to provide the discrete piece of information is shorter than the response generated in the format that is adapted to provide the plurality of pieces of related information.

[4A] The method recited in claim 1, wherein determining the conversation type includes determining a conversational goal corresponding to the natural language utterance,

[4B] wherein the conversation type is determined as a query conversation if the conversational goal is to gather discrete information or perform a discrete task, is determined as a didactic conversation if the conversational goal is to gather related information, or is determined as an exploratory conversation if the natural language utterance has no specific goal.

[5] The method recited in claim 1, wherein the conversation type is determined as a query conversation if the user that spoke the natural language utterance has the leader role, is determined as a didactic conversation if the user that spoke the natural language utterance has the supporter role, or is determined as an exploratory conversation if the user that spoke the natural language utterance has both the leader role and the supporter role.

[6] The method recited in claim 1, wherein determining the conversation type is based on an allocation of information between the user and the speech engine, wherein the conversation type is determined to be a query conversation or a didactic conversation if the user that spoke the natural language utterance has a lower allocation of information related to the natural language utterance than the speech engine, or is determined as a didactic conversation if the user and the speech engine have an equal or alternating allocation of information related to the natural language utterance.

[7] The method recited in claim 1, wherein determining the conversation type further comprises determining an allocation of information between the user and the speech engine.

[8] The method recited in claim 1, wherein generating the response comprises generating, if the conversation type is a query conversation, a response that only provides information requested by the natural language utterance, and generating, if the conversation type is a didactic conversation, a response that controls the conversation and limits the user's future input to a request or query for clarification.

[9] The method recited in claim 1, wherein the natural language utterance is part of a conversation that comprises related natural language utterances from the user, wherein determining whether the user has the leader role or the supporter role comprises determining whether the user controls the conversation or whether the user provides information to the speech engine during the conversation.

[10P] A system for providing a voice interface, comprising:

[10A] a speech engine configured to receive a natural language utterance from a voice-enabled device, the natural language utterance corresponding to a conversation type, wherein the speech engine includes a processor configured to:

[10B] determine the conversation type corresponding to the natural language utterance based on whether a user that spoke the natural language utterance has a leader role in an interaction with the voice-enabled device or has a supporter role in the interaction with the voice-enabled device;

[10C] a response builder configured to generate a response to the natural language utterance with a format based on the conversation type, wherein the

format is adapted to limit the user's future input to interjecting queries or requests for clarification if the user has the supporter role.

[11A] The system recited in claim 10, wherein the speech engine is configured to determine the conversation type further based on a conversational goal corresponding to the natural language utterance, and

[11B] wherein the format of the response is adapted to provide a discrete piece of information to the user, to provide a plurality of pieces of related information to the user, or to obtain one or more pieces of information from the user.

[12A] The system recited in claim 10, wherein the speech engine is configured to determine the conversation type further based on a conversational goal corresponding to the natural language utterance, and

[12B] wherein the conversation type is determined to be a query conversation if the conversational goal is to gather discrete information or perform a discrete task, is determined to be a didactic conversation if the conversational goal is to gather related information, or is determined to be an exploratory conversation if the natural language utterance has no specific goal.

[13] The system recited in claim 10, wherein the speech engine is configured to determine the conversation type based on the role of the user, wherein the conversation type is determined to be a query conversation if the user that spoke the natural language utterance has the leader role, is determined to be a didactic conversation if the user that spoke the natural language utterance has the supporter role, or is determined to be an exploratory conversation if the user that spoke the natural language utterance has both the leader role and the supporter role.

[14] The system recited in claim 10, wherein the speech engine is configured to determine the conversation type based on an allocation of information between the user and the speech engine, wherein the conversation type is determined to be a query conversation or a didactic conversation if the user that spoke the natural language utterance has a lower allocation of information related to the natural language utterance than the speech engine or is determined to be a didactic conversation if the user and the speech engine have an equal or alternating allocation of information.