Exhibit B-Coffman
Invalidity Chart for U.S. Patent No. 8,195,468 in view of U.S. Patent No. 7,137,126 ("Coffman")

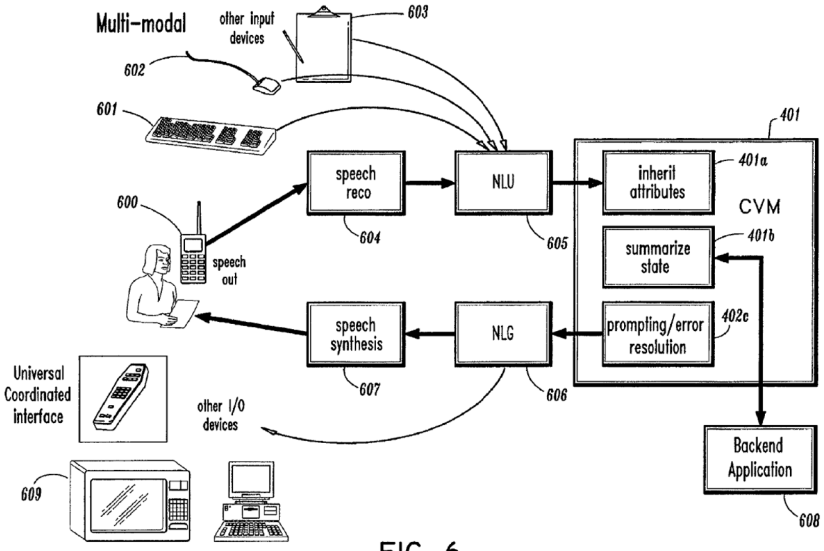
U.S. Patent No. 7,137,126 ("Coffman") was filed Jul. 2, 2001 and the PCT was published Apr. 13, 2000. Coffman qualifies as prior art to U.S. Patent No. 8,195,468 (the "'468 Patent") at least under one or more applicable provisions of 35 U.S.C. §§ 102(a), (b), (e), and/or (g). Coffman discloses and/or renders obvious one or more of claims 19, 25, 26, 27, 28, 29, 30, and 32 of the '468 Patent alone and/or in combination with other references, as set forth in the chart below. To the extent Coffman does not disclose one or more limitations of the claims, such limitations would have been inherent and/or it would have been obvious to combine the disclosures of Coffman with the knowledge of one of ordinary skill in the art and/or with one or more other references, including as identified herein and/or in the Cover Pleading, for the reasons identified herein and/or in the Cover Pleading, to render the claims at-issue in the '468 Patent invalid.

Defendant Bank of America, N.A. ("Bank") incorporates in this chart all applicable qualifications, clarifications, and other statements made in Bank's Invalidity Contentions. This invalidity claim chart is based on Bank's present understanding of claims 19, 25, 26, 27, 28, 29, 30, 32, and Plaintiff Dialect, LLC ("Dialect") apparent construction of the claims, as set forth in Dialect's Infringement Contentions. These invalidity assertions are not an admission that Bank concedes or acquiesces to any claim construction(s) implied or suggested by Dialect the associated infringement claim charts. Nor is Bank asserting any claim construction positions through these charts, including whether the preamble is a limitation. Bank also does not concede or acquiesce that any asserted claim satisfies the requirements of 35 U.S.C. §§ 112 or 101, and submits these invalidity contentions only to the extent Dialect's assertions may be understood.

The excerpts cited herein are exemplary. For any claim limitation, Bank may rely on excerpts cited for any other limitation and/or additional excerpts not set forth fully herein to the extent necessary to provide a more comprehensive explanation for a reference's disclosure of a limitation. Where an excerpt refers to or discusses a figure or figure items, that figure and any additional descriptions of that figure should be understood to be incorporated by reference as if set forth fully herein. To the extent any cell lacks citations to the charted reference(s), this should not be taken as an admission that the reference(s) do(es) not disclose the corresponding limitation but rather indicates that Bank does not presently have sufficient information regarding the charted reference(s) based on Bank's present understanding of the claim limitation.

These invalidity contentions are not an admission by Bank that the accused products or components, including any current or past version of the products or components, are covered by, or infringe the asserted claims, particularly when these claims are properly construed and applied.

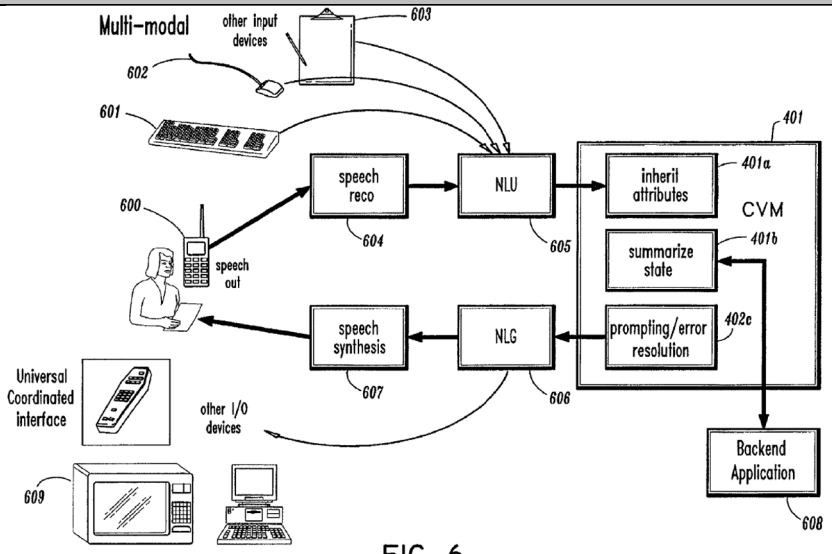
1. Claim 19

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
[p] A method for processing multi-modal natural language inputs, comprising:	To the extent the preamble is limiting, Coffman discloses a "method for processing multi-modal natural language inputs[.]" *E.g.*, Coffman at Fig. 6:  FIG. 6 Coffman at 5:8-30: In another aspect, programming/script languages are utilized that allow the use of any available resources as input or output stream. Using the conversational sub-systems, each input is converted into a binary or ASCII input (lists of attribute-value pairs or n-uples), which can be directly processed by the programming language as built-in objects. Calls, flags and tags are automatically included to transmit between object and processes the conversational meta-information required to correctly interface with the different objects. Indeed, any input in any modality is captured by the

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	dialog manager of the CVM kernel layer as an event that is added to the associated context or context stack. For example, a mouse click or pointer/stylus pointing action followed by the command "I would like to open this" is disambiguated into a set of attribute value pairs: Command: open, Object: Windows or task selected by the last mouse click. Output can be specially formatted according to the needs of the application or user. Multi-modal discourse processing can now be easily built using the new programming tools. In addition, such programming languages and scripts encompasses conversational API between conversational enabled applications and the CVM, as well as CML (conversational markup language). Coffman at 8:33-63: Mixed initiative also implies that the CVM can actively help (take the initiative to help) and coach a user through a task, especially in speech-enable applications, wherein the mixed initiative capability is a natural way of compensating for a display less system or system with limited display capabilities. In general, the CVM complements conventional interfaces and user input/output rather than replacing them. This is the notion of "multi-modality" whereby speech is used in parallel with mouse, keyboard, and other input devices such as a pen. Conventional interfaces can be replaced when device limitations constrain the implementation of certain interfaces. In addition, the ubiquity and uniformity of the resulting interface across devices, tiers and services is an additional mandatory characteristic. It is to be understood that CVM system can to a large extent function with conventional input and/or output media. Indeed, a computer with classical keyboard inputs and pointing devices coupled with traditional monitor display can profit significantly by utilizing the CVM according to the present invention. One example is described in provisional application U.S. Ser. No. 60/128,081, filed on Apr. 7, 1999, entitled "Multi-Modal Shell" which is commonly assigned and incorporated herein by reference (which describes a method for constructing a true multi-modal application with tight synchronization between a GUI modality and a speech modality). In other words, even users who do not want to talk to their computer can also realize a dramatic positive change to their interaction with the CVM enabled machine. Coffman at 14:42-56:

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	The conversational computing system of FIG. 2 further comprises a conversational engine API layer 207 which provides an interface between core engines conversational engines 208 (e.g., speech recognition, NL parsing, NLU, TTS and speech compression/decompression engines) and the applications using them. The engine API layer 207 also provides the protocols to communicate with core engines whether they be local or remote. An I/O API layer 210 provides an interface with conventional I/O resources 211 such as a keyboard, mouse, touch screen, keypad, etc. (for providing a multi-modal conversational UI) and an audio subsystem for capturing speech I/O (audio in/audio out). The I/O API layer 210 provides device abstractions, I/O abstractions and UI abstractions. The I/O resources 211 will register with the CVM kernel layer 202 via the I/O API layer 210. Coffman at 27:65-28:38: Referring now to FIG. 6, a diagram illustrates a conversational input/output interface in accordance with one embodiment of the present invention. As illustrated, a conversational input interface according to an embodiment of the present invention can process multi-modal input, that is, files/streams/resources, speech via a phone 600, keyboard 601, pointing devices 602, handwriting devices 603, including natural interfaces. This means that all the input and output events across all the modalities are caught and transferred to the dialog manager (that also stores it appropriately in the context stack). Spoken input from a speech client (e.g., telephone 600) is subject to a speech recognition process 604 and other input (e.g., keyboard, mouse clicks etc) are subject to NLU processing 605. Each input is subject to attribute acquisition (401 a) whereby the attribute value n-uples are acquired from the input. A summarization process 401 b is performed whereby the attribute value n-uples are added to the context and then verifies with the syntax of the back-end application 608 whether the query is complete, incomplete, or ambiguous. The backend accesses are also tracked by the dialog manager and the context manager. It is sometimes possible to distribute some of the "intelligence" tot he backend by loading some disambiguation capabilities (a feature of the dialog manager) to the backend. Individually, each input stream behaves the conventionally. The key conversational aspect is in the input procedure wherein commands can be entered in NLU (to provide natural language understanding of input queries) or

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	in FSG mode (for constrained input according to rules: grammar and vocabulary, as opposed to free natural input). Commands or queries can be completed or corrected by completing missing fields or by correcting incorrect fields for the active task. As such, the CVM introduces new issues not met with conventional OS: simultaneous input streams to be merged which create input ambiguity. For example, input may now combine input keyed on the keyboard, handwritten input and speech input, not to mention possible input from re-directed streams. Therefore, the present invention provides a mechanism to resolve any ambiguity. This may be performed as explained in the above-incorporated application U.S. Ser. No. 60/128,081. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[a] receiving a multi-modal natural language input at a conversational voice user interface, the multi-modal input including a natural language utterance and a non-speech input provided by a user, wherein a transcription module coupled to the conversational voice user interface transcribes the non-	Coffman discloses "receiving a multi-modal natural language input at a conversational voice user interface, the multi-modal input including a natural language utterance and a non-speech input provided by a user, wherein a transcription module coupled to the conversational voice user interface transcribes the non-speech input to create a non-speech-based transcription[.]" *E.g.*, Coffman at Fig. 6:

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
speech input to create a non-speech-based transcription;	 FIG. 6 Coffman at 5:8-30: In another aspect, programming/script languages are utilized that allow the use of any available resources as input or output stream. Using the conversational sub-systems, each input is converted into a binary or ASCII input (lists of attribute-value pairs or n-uples), which can be directly processed by the programming language as built-in objects. Calls, flags and tags are automatically included to transmit between object and processes the conversational meta-information required to correctly interface with the different objects. Indeed, any input in any modality is captured by the dialog manager of the CVM kernel layer as an event that is added to the associated context or context stack. For example, a mouse click or pointer/stylus pointing action followed by the command "I would like to open this" is disambiguated into a set of attribute value pairs: Command: open, Object: Windows or task selected by the last mouse click. Output can be specially formatted according to the needs of the application or user. Multi-modal discourse processing can now be easily built using the new programming tools. In addition, such

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	programming languages and scripts encompasses conversational API between conversational enabled applications and the CVM, as well as CML (conversational markup language). Coffman at 8:33-63: Mixed initiative also implies that the CVM can actively help (take the initiative to help) and coach a user through a task, especially in speech-enable applications, wherein the mixed initiative capability is a natural way of compensating for a display less system or system with limited display capabilities. In general, the CVM complements conventional interfaces and user input/output rather than replacing them. This is the notion of "multi-modality" whereby speech is used in parallel with mouse, keyboard, and other input devices such as a pen. Conventional interfaces can be replaced when device limitations constrain the implementation of certain interfaces. In addition, the ubiquity and uniformity of the resulting interface across devices, tiers and services is an additional mandatory characteristic. It is to be understood that CVM system can to a large extent function with conventional input and/or output media. Indeed, a computer with classical keyboard inputs and pointing devices coupled with traditional monitor display can profit significantly by utilizing the CVM according to the present invention. One example is described in provisional application U.S. Ser. No. 60/128,081, filed on Apr. 7, 1999, entitled "Multi-Modal Shell" which is commonly assigned and incorporated herein by reference (which describes a method for constructing a true multi-modal application with tight synchronization between a GUI modality and a speech modality). In other words, even users who do not want to talk to their computer can also realize a dramatic positive change to their interaction with the CVM enabled machine. Coffman at 14:42-56: The conversational computing system of FIG. 2 further comprises a conversational engine API layer 207 which provides an interface between core engines conversational engines 208 (e.g., speech recognition, NL parsing, NLU, TTS and speech compression/decompression engines) and the applications using them. The engine API layer 207 also provides the protocols to communicate with core engines whether they be local or remote. An I/O API layer 210 provides an interface

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	with conventional I/O resources 211 such as a keyboard, mouse, touch screen, keypad, etc. (for providing a multi-modal conversational UI) and an audio subsystem for capturing speech I/O (audio in/audio out). The I/O API layer 210 provides device abstractions, I/O abstractions and UI abstractions. The I/O resources 211 will register with the CVM kernel layer 202 via the I/O API layer 210. Coffman at 27:65-28:38: Referring now to FIG. 6, a diagram illustrates a conversational input/output interface in accordance with one embodiment of the present invention. As illustrated, a conversational input interface according to an embodiment of the present invention can process multi-modal input, that is, files/streams/resources, speech via a phone 600, keyboard 601, pointing devices 602, handwriting devices 603, including natural interfaces. This means that all the input and output events across all the modalities are caught and transferred to the dialog manager (that also stores it appropriately in the context stack). Spoken input from a speech client (e.g., telephone 600) is subject to a speech recognition process 604 and other input (e.g., keyboard, mouse clicks etc) are subject to NLU processing 605. Each input is subject to attribute acquisition (401 a) whereby the attribute value n-uples are acquired from the input. A summarization process 401 b is performed whereby the attribute value n-uples are added to the context and then verifies with the syntax of the back-end application 608 whether the query is complete, incomplete, or ambiguous. The backend accesses are also tracked by the dialog manager and the context manager. It is sometimes possible to distribute some of the "intelligence" tot he backend by loading some disambiguation capabilities (a feature of the dialog manager) to the backend. Individually, each input stream behaves the conventionally. The key conversational aspect is in the input procedure wherein commands can be entered in NLU (to provide natural language understanding of input queries) or in FSG mode (for constrained input according to rules: grammar and vocabulary, as opposed to free natural input). Commands or queries can be completed or corrected by completing missing fields or by correcting incorrect fields for the active task. As such, the CVM introduces new issues not met with conventional OS: simultaneous input streams to be merged which create input ambiguity. For example, input may now combine input keyed on the keyboard, handwritten input and speech input, not to mention possible input from re-directed streams. Therefore, the present

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	invention provides a mechanism to resolve any ambiguity. This may be performed as explained in the above-incorporated application U.S. Ser. No. 60/128,081. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[b] identifying the user that provided the multi-modal input;	Coffman discloses "identifying the user that provided the multi-modal input[.]" *E.g.*, Coffman at 27:51-64: It is to be appreciated that conversational biometrics can be used to collect any context and meta information on the user not only to customize or adapt for purposes of user preferences or to authorize a query, but also to use the information to perform more robust recognition. Accordingly, any information can be accumulated to recognize the user. Namely, the usual phrasing of a query, the type of query phrased, command frequency (often used, not often used), preferred applications, time or usage, etc. Conversational biometrics may be built using the methods disclosed in U.S. Pat. No. 5,897,616 entitled "Apparatus and Methods for Speaker Verification/Identification/Classification Employing Non-Acoustic and/or Acoustic Models, which is commonly assigned and incorporated herein by reference. Coffman at 21:45-50: The conversational system 400 further comprises a plurality of conversational resource subsystems (engines) 407 including, for example, a speech recognition system 408, a speaker

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	recognition system 409, a natural language understanding and natural language parsing system 410 and a text-to-speech synthesis (TTS) system 411. Coffman at 35:44-53: Conversational VM can adapt dynamically to the preferences of multiple users and to the active context. It allows multiple users while actively running. In a speech-based system, each command can be used to perform text-independent speaker identification. Any change of user automatically implies the creation of a new active context which pushes the previous context down the context stack, unless the new active context is waived explicitly by the new user or the active application. User changes automatically change the priority along the context stack to first handle a task associated to the active user. Coffman at 34:31-35:3: Another feature provided by a conversational CVM system is "conversational security," whereby meta-information relating to the author and/or modifier of local or remote files, especially executable files, can be used for security purposes. In particular, with speech-based conversational systems, since each command conveys not only the formulation of the query but also enough information for authentication of the user using, text-independent speaker verification can be used to identify and verify a user. In this manner, the automatic (and transparent) authentication of the user can be made whenever a query to a restricted resource is made, based on security meta-information associated to the resource. As noted above, all the information collected about the user queries and history can be used to contribute to the recognition (ID or verification) of the user. The authentication an be performed either directly on the request or using non-expired information acquired shortly before the query. In particular, authorization for access to files or application can on a query by query basis. For instance, if a user requests a restricted service, the request may be verified with respect to the set of users that are pre-authorize to access that specific service. The authentication can be performed via open-set speaker identification performed on the request (e.g.,

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	file access, directory access, application opening, executables, connections, encryption/decryption, digital certification/signature). Resources having different passwords or a user ID associated with a similar user can be seamlessly accessed with no explicit login or password authentication. In any event, non-obtrusive user authentication can be continuously and transparently performed through user dialog. In accordance with the idea that a conversational VM can be implemented even with no speech input, the stack of contexts should contain the identity of the user as the most recently authenticated identity. In addition, each resource should contain the list of authorized users as well as some security requirements (e.g. in a non-speech case the expiration date of the latest authentication). Of course key-strokes or pen based authentication can also be considered, but it is not at all mandatory. Coffman at 42:17-32: You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. Coffman at 11:27-45: Low-Level Dialog Conversational Foundation Classes: (multi-modal feature available where appropriate) (with CVM handle when distributed)

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	1. Select an item from list 2. Field filing with grammar 3. Acoustic Enroll speaker 4. Acoustic Identify speaker 5. Acoustic Verify speaker 6. Verify utterance 7. Add to list 8. Enroll utterance 9. Get input from NL 10. Disambiguate *E.g.*, Coffman at 12:13-43: Intermediate-Level Dialog Conversational Foundation Classes (multi-modal feature available where appropriate) (with CVM handle when distributed) 1. Form_filling 2. Request_confirmation 3. Identify user by dialog 4. Enrol_user by dialog 5. Speech_biometrics_identify 6. Verify_user by dialog 7. Correct_input 8. Speech_biometrics_identify 9. Speech_biometrics_verify 10. Speech_biometrics_enrol 11. Manage_table 12. Fill_free_field 13. Listen_to_TTS 14. Listen_to_playback 15. Simulltaneous_form_filling

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	16. Simultaneous_classes_dialog 17. Summarize_dialog To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[c] creating a speech-based transcription of the natural language utterance using a speech recognition engine and a semantic knowledge-based model, wherein the semantic knowledge-based model includes a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface, a general cognitive model derived from one or more prior interactions between a plurality of users and the conversational voice user interface, and an	Coffman discloses "creating a speech-based transcription of the natural language utterance using a speech recognition engine and a semantic knowledge-based model, wherein the semantic knowledge-based model includes a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface, a general cognitive model derived from one or more prior interactions between a plurality of users and the conversational voice user interface, and an environmental model derived from an environment of the identified user and the conversational voice user interface[.]" *E.g.*, Coffman at 4:47-64: In another aspect, the CVM provides the capability to have natural dialog with NLU, NLG, contexts and mixed-initiatives sorted across multiple tasks, processes and discourses (with multiple domains). A conversational input interface is provided whereby a set of multi-mode input streams are each transcribed into an ASCII command or query (i.e., lists of attribute-value pairs or n-uples). Each input entity (command, NLU query field or argument unit (isolated letter, word, etc.) is associated with time-marks and appended accordingly to a compounded input stream. Two or more stream having the same time-marks are prioritized based on when each input stream contributed previously or the priority that each application/input stream received on the basis of

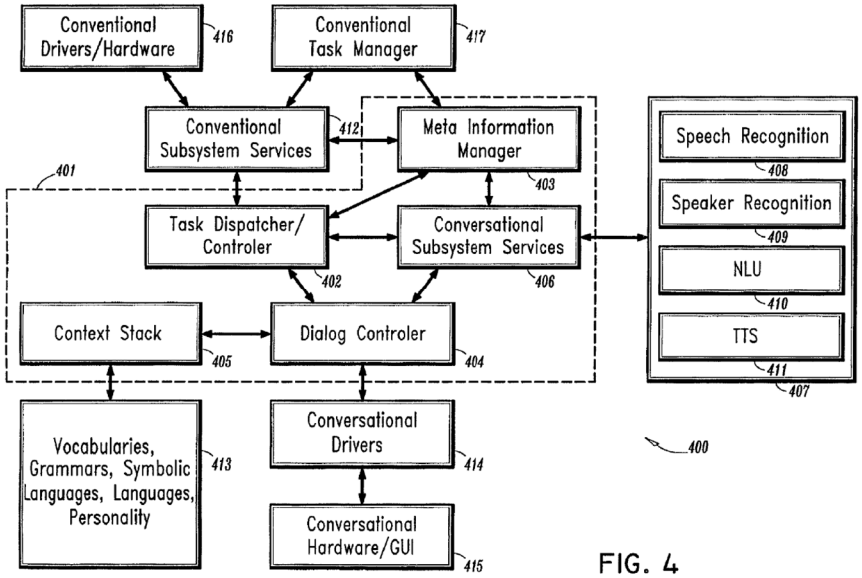
Claim 19 of the '468 Patent	Exemplary Citations to Coffman
environmental model derived from an environment of the identified user and the conversational voice user interface;	the context history. Compounded inputs are checked against possible FSG and dictionaries and optionally fed back to the user. Each resource exchanges their conversational capabilities and the input stream is tailored to only exchange relevant information. Coffman at 28:39-57: In accordance with the present invention, the input problem is treated as a merge of the output of multiple decoders, ASCII transcription or a list of attribute value n-uples. Each input stream is converted into its ASCII transcription and aligned with input time-marks via speech recognition processing 604. When different input stream are associated to the same task, the transcripts are merged as follows. First, commands and queries are sorted based on the time marks and appended to a single data stream. Command formulation can be checked against FSG rules and re-sorted to satisfy the grammar rules. NLU queries do not necessarily require re-sorting. For NLU queries, the symbolic fields are filled for each stream, then compounded at the level of the final input stream. Arguments such as spelling and alpha-numeric code do not exploit grammar rules or NLU to solve ordering ambiguity. Time-marks are used similarly to build a unique stream. However, the input is fed back to the user for confirmation, with possible pre-filtering using a dictionary or FSG rule book which is application-dependent. Coffman at 30:36-50: The task dispatcher/controller 402 divides each command/process into multiple actions, starts the associated threads/processes with the appropriate priority and relates/inserts them within the context stack 405. The task dispatcher 402 allocates each resource and shares them between the different spawned actions, and controls handles and streams to and from the resources. Based on the modality (pointer, keyboard, file, speech), the task dispatcher 402 redirects the stream to the appropriate conversational subsystems or conventional subsystem with speech inputs being transcribed/understood. The output of these subsystems is run down the context stack 405 to extract the active query and complete it. On the other hand, outputs are queued based on the priority levels of each task and dispatched sequentially to the output resource.

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	Coffman at 6:56-67: In yet another aspect, conversational help, manuals and support is provided through a ubiquitous coordinated conversational interface, using local and remote resources, usage history of a user and agents to complete request, guide through procedure, search for information and upgrade/install new applications. In addition, help information can be accessed using NLU queries to access the help information or on the basis of the meta-information associated to the current user (history) and on the basis of the arguments that are missing or modified using mixed initiative. The dialog provided by each application is tuned to the preferences or level of expertise of the user. Coffman at 16:29-33: The global history stores, for example, the information that is associated with all the applications and actions taking during a conversational session (i.e., the history of the dialog between user and machine for a current session (or from when the machine was activated). Coffman at 15:50-16:17: The task dispatcher/manager 221 dispatches and coordinates different tasks and processes that are spawned (by the user and machine) on local and networked conventional and conversational resources (explained in further detail below). The meta information manager 222 manages the meta-information associated with the system via a meta-information repository 218. The meta information manager 218 and repository 218 collect all the information typically assumed known in a conversational interaction but not available at the level of the current conversation. Examples are: a-priori knowledge: cultural, educational assumptions and persistent information: past request, references, information about the user, the application, news, etc. It is typically the information that needs to be preserved and persist beyond the length/life of the conversational history/context and the information that is expected to be common knowledge for the conversation and therefore, has never been defined during the current and possible past conversational interactions. Also, as described below, shortcuts to commands, resources and macros, etc. are managed by the meta-information manager 222 and stored in the meta information repository 218. In addition, the meta-

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	information repository 21 includes a user-usage log based on user identity. It is to be appreciated that services such as conversational help and assistance, as well as some dialog prompts (introduction, questions, feedback etc) provided by the CVM system can be tailored based on the usage history of the user as stored in the meta-information repository 218 and associated with the application. If a user has been previously interacting with a given application, an explanation can be reduced assuming that it is familiar to the user. Similarly, if a user commits many errors, the explanations can be more complex, as multiple errors is interpreted as user uncertainty, unfamiliarity, or incomprehension/misunderstanding of the application or function. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[d] merging the speech-based transcription and the non-speech-based transcription to create a merged transcription;	Coffman discloses "merging the speech-based transcription and the non-speech-based transcription to create a merged transcription[.]" *E.g.*, Coffman at 4:47-56: In another aspect, the CVM provides the capability to have natural dialog with NLU, NLG, contexts and mixed-initiatives sorted across multiple tasks, processes and discourses (with multiple domains). A conversational input interface is provided whereby a set of multi-mode input streams are each transcribed into an ASCII command or query (i.e., lists of attribute-value pairs or n-uples). Each input entity (command, NLU query field or argument unit (isolated letter, word, etc.) is associated with time-marks and appended accordingly to a compounded input stream.

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	Coffman at 27:65-28:38: Referring now to FIG. 6, a diagram illustrates a conversational input/output interface in accordance with one embodiment of the present invention. As illustrated, a conversational input interface according to an embodiment of the present invention can process multi-modal input, that is, files/streams/resources, speech via a phone 600, keyboard 601, pointing devices 602, handwriting devices 603, including natural interfaces. This means that all the input and output events across all the modalities are caught and transferred to the dialog manager (that also stores it appropriately in the context stack). Spoken input from a speech client (e.g., telephone 600) is subject to a speech recognition process 604 and other input (e.g., keyboard, mouse clicks etc) are subject to NLU processing 605. Each input is subject to attribute acquisition (401 a) whereby the attribute value n-uples are acquired from the input. A summarization process 401 b is performed whereby the attribute value n-uples are added to the context and then verifies with the syntax of the back-end application 608 whether the query is complete, incomplete, or ambiguous. The backend accesses are also tracked by the dialog manager and the context manager. It is sometimes possible to distribute some of the "intelligence" tot he backend by loading some disambiguation capabilities (a feature of the dialog manager) to the backend. Individually, each input stream behaves the conventionally. The key conversational aspect is in the input procedure wherein commands can be entered in NLU (to provide natural language understanding of input queries) or in FSG mode (for constrained input according to rules: grammar and vocabulary, as opposed to free natural input). Commands or queries can be completed or corrected by completing missing fields or by correcting incorrect fields for the active task. As such, the CVM introduces new issues not met with conventional OS: simultaneous input streams to be merged which create input ambiguity. For example, input may now combine input keyed on the keyboard, handwritten input and speech input, not to mention possible input from re-directed streams. Therefore, the present invention provides a mechanism to resolve any ambiguity. This may be performed as explained in the above-incorporated application U.S. Ser. No. 60/128,081. Coffman at 28:39-57:

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	In accordance with the present invention, the input problem is treated as a merge of the output of multiple decoders, ASCII transcription or a list of attribute value n-uples. Each input stream is converted into its ASCII transcription and aligned with input time-marks via speech recognition processing 604. When different input stream are associated to the same task, the transcripts are merged as follows. First, commands and queries are sorted based on the time marks and appended to a single data stream. Command formulation can be checked against FSG rules and re-sorted to satisfy the grammar rules. NLU queries do not necessarily require re-sorting. For NLU queries, the symbolic fields are filled for each stream, then compounded at the level of the final input stream. Arguments such as spelling and alpha-numeric code do not exploit grammar rules or NLU to solve ordering ambiguity. Time-marks are used similarly to build a unique stream. However, the input is fed back to the user for confirmation, with possible pre-filtering using a dictionary or FSG rule book which is application-dependent. Coffman at 29:1-13: In summary, the input procedure provides a set of multi-mode input streams, each transcribed into an ASCII command, query, or list of attribute value n-uples. Each input entity (command, NLU query field or argument unit (isolated letter, word etc.) are associated to time-marks and appended accordingly to a compounded input stream. Should two or more stream have exactly the same time-marks, they are prioritized based on when each input stream contributed previously. Compounded inputs are checked against possible FSG and dictionaries and optionally fed back to the user. Each resource exchanges their conversational capabilities and the input stream is tailored to only exchange relevant information. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[e] identifying one or more entries in a context stack matching information contained in the merged transcription;	Coffman discloses "identifying one or more entries in a context stack matching information contained in the merged transcription[.]" *E.g.*, Coffman at Fig. 4:  FIG. 4 Coffman at 4:6-26: In another aspect, a CVM kernel layer (or CVM controller) comprises a meta-information manager, a resource manager, a context stack, a global history, a dialog manager and a task

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	dispatcher, for managing the dialog and selecting the active dialog, context, and application. The context stack accumulates the context (full query arguments and state/mode—i.e. query arguments already introduced, any I/O event, and event produced by an application) of each active process with an activated discourse along with any data needed for input understanding (e.g. active FSG, topic, vocabulary or possible queries for a speech input). The CVM kernel coordinates the different tasks and processes that are spawned on local and networked conventional and conversational resources. The CVM kernel layer keeps track of these resources, transmit input to the appropriate conversational subsystems and arbitrate between devices, state and applications. The CVM kernel layer also coordinates the output generation and prioritization according to the active conversation and conversation history, delayed returns, delegation across network resources and task delegation and memorization. Coffman at 22:5-46: The core of the CVM 401 is the context stack 405 which operates and is managed under the control of the dialog controller 404 (it is to be understood that the context stack 405 id directly related to the global history and meta information repository discussed above). In general, the context stack 405 accumulates the context (i.e., full query arguments list of attribute value n-uples, and state/mode) of each active process with an activated discourse (i.e., conversational interaction associated with a given task/process/thread) along with any data files 413 (or at least identifiers of such conversational arguments) for the different engines that may be needed for input understanding (e.g., files or arguments that the engines use for performing their respective tasks such as active FSG, topic, vocabulary, HMM (hidden markov models), voiceprints, language models or possible queries for a speech input). In other words, the term "context" refers to the state of each discourse (whether active or normative), which keeps track of the past history of the discourse, its current state, and the specific characteristics and full query arguments of the corresponding task (e.g, vocabulary file, language model, parsing, tags, voiceprint, TTS rules, grammar, NLU etc. of each active task/process) along with any data needed for input understanding. It is to be appreciated that, in a distributed CVM (as explained below), the context stack (as well as other CVM components) may be directly associated with networked services

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	(i.e., distributed over the network) (as described above with respect to the context and global history). More specifically, each new task, process, or thread creates a new stack entry and is associated with a discourse. Each application may be associated with multiple discourses (e.g. the application management discourse and the application content navigation discourses). Each context associated with a given discourse comprises the latest requests made to the corresponding process/task/thread as well as the latest output. The context of a given discourse is also associated with, e.g., any active grammars, vocabularies and symbolic language which maps the actual query. Again, the latest information is stored in the history and context stacks. Past history and context and other information is managed by the meta information manager 403 and stored as part of the meta information. Coffman at 31:23-49: Completion/search for the active context is performed from context to context down the stack. That is, new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user. As soon a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active. Until the active command is completed, or until a new command is provided, the selected context is marked active, and pushed to the top of the context stack 405. When a message is returned to the user, the context is updated and then pushed to the top of the context stack 405 under the active context. The active context is updated to inform of the existence of a returned value. This can also be done at the level of the superseding CVM session discourse, which can be in the stack or always besides the stack and then searched right after the active context, before going down the stack. Simultaneously completed tasks result in contexts that are arranged under the active context according to CVM priorities (e.g. FIFO or FILO). Active contexts sufficiently completed to generate a task will be pushed down the stack under the next or all the returned contexts. Or it could become the active discourse. This may be done automatically or when commanded by the user. This stack structure

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	allows to maintain non-ambiguous conversational interactions with multiple tasks, threads or processes. Coffman at 42:58-43:6: With CVM running on the desktop and on the PVA and CVM aware smart phone and PDA, the application developer must only hook to the CVM API. It involves registering all its conversational capabilities and requirements: 1. Active vocabulary, finite state grammar and language models to control the application; 2 Symbolic mapping if NLU is supported or at list a context state list; 3 Associated relevant meta-information/categories in particular to allow categorization of the for the output; 4 Conversational I/O information: does it directly control the input/output or is it a silent partner, client to a conversational I/O provider; and 5 CVM capabilities/state: does it run a CVM; is it a CVM client; is it a master, slave or partner CVM. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[f] determining a most likely context for the multi-modal	Coffman discloses "determining a most likely context for the multi-modal input based on the identified entries[.]"

Claim 19 of the '468 Patent	**Exemplary Citations to Coffman**
input based on the identified entries;	*E.g.*, Coffman at 31:23-30: Completion/search for the active context is performed from context to context down the stack. That is, new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user. As soon a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[g] identifying a domain agent associated with the most likely context for the multi-modal input;	Coffman discloses "identifying a domain agent associated with the most likely context for the multi-modal input[.]" *E.g.*, Coffman at 32:31-47: When context stacks are added, any overlap can be removed at the bottom of the stack. The user, platform or application can decide to only preserve save portion of the stack. Conversational assistants perform such tasks. They can be implemented by agents and daemons simply running on their own and re-interact with the user only when producing output. Their output is sent to the user according to the priority level of the task. When becoming active the user can easily update the task associated to the agent. Conversational memorization, are rather tasks inserted at the bottom of the stack and executed only when the stack is emptied at the end of the session. Occasionally, they can be inserted higher in the stack or pushed to the top of the stack at a pre-decided moment.

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	Memorization tasks are executed only when active. The memorization feature affords the capability to memorize past actions, preferences and instructions. Coffman at 42:1-32: When multiple devices are conversationally connected, they will coordinate their interfaces so that all the devices can be controlled through the universal CUI. This concept may be illustrated by the following example. Assume that you are driving home one night and remember that your spouse asked you to buy some goods at a new grocery store. After finding the message on your answering machine, you rapidly transcribed it into a memo on your desktop using a speech recognition software. However, you forgot to print it or transfer it on your PDA. It does not matter if your desktop PC runs a CVM since you have, in your car, a conversational PDA, a conversational car PC (PVA, Personal Vehicle Assistant) and a conversational smart phone. Further assume that the PVA runs an embedded CVM while the two other applications are conversationally aware, i.e., you can control them through the CVM running on the PVA. You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[h] communicating a request to the identified domain agent; and	Coffman discloses "communicating a request to the identified domain agent[.]" *E.g.*, Coffman at 42:1-32: When multiple devices are conversationally connected, they will coordinate their interfaces so that all the devices can be controlled through the universal CUI. This concept may be illustrated by the following example. Assume that you are driving home one night and remember that your spouse asked you to buy some goods at a new grocery store. After finding the message on your answering machine, you rapidly transcribed it into a memo on your desktop using a speech recognition software. However, you forgot to print it or transfer it on your PDA. It does not matter if your desktop PC runs a CVM since you have, in your car, a conversational PDA, a conversational car PC (PVA, Personal Vehicle Assistant) and a conversational smart phone. Further assume that the PVA runs an embedded CVM while the two other applications are conversationally aware, i.e., you can control them through the CVM running on the PVA. You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use.

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.
[i] generating a response to the user from content provided by the identified domain agent as a result of processing the request.	Coffman discloses "generating a response to the user from content provided by the identified domain agent as a result of processing the request[.]" *E.g.*, Coffman at 42:17-32: You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover

Claim 19 of the '468 Patent	Exemplary Citations to Coffman
	Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

2. Claim 25

Claim 25 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 19, further comprising verifying an identity of the user based on voiceprint matching, password matching, or pass-phrase matching.	Coffman discloses "verifying an identity of the user based on voiceprint matching, password matching, or pass-phrase matching[.]" *E.g.*, Coffman at 27:51-64: It is to be appreciated that conversational biometrics can be used to collect any context and meta information on the user not only to customize or adapt for purposes of user preferences or to authorize a query, but also to use the information to perform more robust recognition. Accordingly, any information can be accumulated to recognize the user. Namely, the usual phrasing of a query, the type of query phrased, command frequency (often used, not often used), preferred applications, time or usage, etc. Conversational biometrics may be built using the methods disclosed in U.S. Pat. No. 5,897,616 entitled "Apparatus and Methods for Speaker Verification/Identification/Classification Employing Non-Acoustic and/or Acoustic Models, which is commonly assigned and incorporated herein by reference. Coffman at 21:45-50: The conversational system 400 further comprises a plurality of conversational resource subsystems (engines) 407 including, for example, a speech recognition system 408, a speaker

Claim 25 of the '468 Patent	Exemplary Citations to Coffman
	recognition system 409, a natural language understanding and natural language parsing system 410 and a text-to-speech synthesis (TTS) system 411. Coffman at 35:44-53: Conversational VM can adapt dynamically to the preferences of multiple users and to the active context. It allows multiple users while actively running. In a speech-based system, each command can be used to perform text-independent speaker identification. Any change of user automatically implies the creation of a new active context which pushes the previous context down the context stack, unless the new active context is waived explicitly by the new user or the active application. User changes automatically change the priority along the context stack to first handle a task associated to the active user. Coffman at 34:31-35:3: Another feature provided by a conversational CVM system is "conversational security," whereby meta-information relating to the author and/or modifier of local or remote files, especially executable files, can be used for security purposes. In particular, with speech-based conversational systems, since each command conveys not only the formulation of the query but also enough information for authentication of the user using, text-independent speaker verification can be used to identify and verify a user. In this manner, the automatic (and transparent) authentication of the user can be made whenever a query to a restricted resource is made, based on security meta-information associated to the resource. As noted above, all the information collected about the user queries and history can be used to contribute to the recognition (ID or verification) of the user. The authentication an be performed either directly on the request or using non-expired information acquired shortly before the query. In particular, authorization for access to files or application can on a query by query basis. For instance, if a user requests a restricted service, the request may be verified with respect to the set of users that are pre-authorize to access that specific service. The authentication can be performed via open-set speaker identification performed on the request (e.g.,

Claim 25 of the '468 Patent	Exemplary Citations to Coffman
	file access, directory access, application opening, executables, connections, encryption/decryption, digital certification/signature). Resources having different passwords or a user ID associated with a similar user can be seamlessly accessed with no explicit login or password authentication. In any event, non-obtrusive user authentication can be continuously and transparently performed through user dialog. In accordance with the idea that a conversational VM can be implemented even with no speech input, the stack of contexts should contain the identity of the user as the most recently authenticated identity. In addition, each resource should contain the list of authorized users as well as some security requirements (e.g. in a non-speech case the expiration date of the latest authentication). Of course key-strokes or pen based authentication can also be considered, but it is not at all mandatory. Coffman at 42:17-32: You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. Coffman at 11:27-45: Low-Level Dialog Conversational Foundation Classes: (multi-modal feature available where appropriate) (with CVM handle when distributed)

Claim 25 of the '468 Patent	Exemplary Citations to Coffman
	1. Select an item from list 2. Field filing with grammar 3. Acoustic Enroll speaker 4. Acoustic Identify speaker 5. Acoustic Verify speaker 6. Verify utterance 7. Add to list 8. Enroll utterance 9. Get input from NL 10. Disambiguate *E.g.*, Coffman at 12:13-43: Intermediate-Level Dialog Conversational Foundation Classes (multi-modal feature available where appropriate) (with CVM handle when distributed) 1. Form_filling 2. Request_confirmation 3. Identify user by dialog 4. Enrol_user by dialog 5. Speech_biometrics_identify 6. Verify_user by dialog 7. Correct_input 8. Speech_biometrics_identify 9. Speech_biometrics_verify 10. Speech_biometrics_enrol 11. Manage_table 12. Fill_free_field 13. Listen_to_TTS 14. Listen_to_playback 15. Simulltaneous_form_filling

Claim 25 of the '468 Patent	Exemplary Citations to Coffman
	16. Simultaneous_classes_dialog 17. Summarize_dialog To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

3. Claim 26

Claim 26 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 19, further comprising receiving one or more follow-up multi-modal inputs at the conversational voice user interface, the follow-up multi-modal inputs including at least one of a follow-up natural language utterance or a follow-up non-speech input.	Coffman discloses "receiving one or more follow-up multi-modal inputs at the conversational voice user interface, the follow-up multi-modal inputs including at least one of a follow-up natural language utterance or a follow-up non-speech input[.]" *E.g.*, Coffman at 42:17-32: You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA

Claim 26 of the '468 Patent	Exemplary Citations to Coffman
	handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. Coffman at 42:33-39: You now instruct the PVA to guide you turn by turn to the store. Your position is computed, the location of the store is fetched, locally or from a server, and an itinerary is computed to take into account the latest traffic information. At any time, you can request navigation information about where you are, what to do next, how far to go or even request a different itinerary. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

4. Claim 27

Claim 27 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 26, wherein the identified domain agent updates the context stack and the semantic knowledge-	Coffman discloses "wherein the identified domain agent updates the context stack and the semantic knowledge-based model in response to processing the request[.]" *E.g.*, Coffman at 30:10-19:

Claim 27 of the '468 Patent	**Exemplary Citations to Coffman**
based model in response to processing the request.	By way of example, the task dispatcher/controller 402 will decide on the basis of the context stack 405 the different statistical parsers that must operate on a query for the dialog controller 404 to identify the active context and complete the query. These actions must be appropriately prioritized so as to, e.g., execute completed queries and update the context stack 405, provide feedback to the user for incomplete or ambiguous queries/command, allow new input to be decoded and run down the context stack 405, and return output of executed or running processes. Coffman at 31:8-13: On the other hand, complete and non-ambiguous commands will result in certain results (e.g., outputs or actions). These results are similarly returned to the appropriate conversational engine 407 with a priority level and update the context, unless if re-directed by the user as in conventional systems. Coffman at 31:23-37: Completion/search for the active context is performed from context to context down the stack. That is, new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user. As soon a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active. Until the active command is completed, or until a new command is provided, the selected context is marked active, and pushed to the top of the context stack 405. When a message is returned to the user, the context is updated and then pushed to the top of the context stack 405 under the active context. The active context is updated to inform of the existence of a returned value. Coffman at 31:50-67: If the request is complete, it will be executed, pending possible request for confirmation by the user, e.g. when it is irreversible. Otherwise, mixed initiative is used to continue the completion or correct the query/command. Whenever, a command/request progresses, option is opened in the

Claim 27 of the '468 Patent	Exemplary Citations to Coffman
	context for rejection of the discourse by the user. This would mean, restoring the previous stack status (and program status) and pursuing down the stack. The user would have to explicitly request going back up the stack. If the user rejects or immediately completes his/her input prior to execution or notification of execution to the user, the new input is appended to the active utterances and the search is re-started from the top of the stack. Any other utterance provided by the user, before the active context is established, is stored in a buffer and considered as appended to the active utterance (speech utterance or any other mode of input). The context stack is updated pending on voice, keyboard, mouse or any other input or command and on the application output. Coffman at 32:22-28: The context stack 405 are rearranged at the time the memorized task interacts with the user. The current context stack is stored and the old context stack is added on top of the stack, with possible updates as programmed by the user or application developer or imposed by CVM, based on intermediate changes die to the evolution of the context and dialogs between launching the task and its completion. Coffman at 36:4-26: In summary, with respect to conversational multi-users and conversational security, it is to be appreciated that dialogs, categories, meta-information, and access to resources can be a function of the identity of the user and its associated meta-information history. And conversely, the conversational information collected on a query can be used to recognize the user. The meta-information associated with each object can be consulted and updated before and after each action or access. When an object is created, modified or consulted, information about the user is added to its meta-information so that the meta-information comprises security and preference fields associated to each object. Access to an object is based on its content, date of creation, history of access and modification and other meta-information. Access is controlled or configured not only based on the identity of the user but on additional meta-information like the date, the usage history, the opened applications etc. In other words, it is possible to allow a person to access a file provided that the file is opened to display on the screen or play back or execution. However, the

Claim 27 of the '468 Patent	Exemplary Citations to Coffman
	person is denied access to open the file to copy its content to another object. In addition, meta-information can be tagged in an un-erasable fashion to an object. Coffman at 37:52-38:7: In all these cases, the menu used for navigation by selection through the web pages or the file system, or whatever other hierarchical structure of object and resources can be appropriately presented in one of various complementary manners. For instance, at the moment of registration of a networked object, the menu can carry meta-information about its structure. Moreover, the system can locally keep track in the meta-information that it associates to each object of the structure (skeleton) of the structure (conversational structure skeletons are described in detail in the patent application Ser. No. 09/806,549, filed concurrently herewith, entitled "Structure Skeletons For Efficient Voice Navigation Through Generic Hierarchical Objects", which is commonly assigned and incorporated herein by reference. Moreover, the system can periodically update its skeleton information, during off-peak use of the CPU. The system can periodically spider any local or external resource and hierarchical object. Alternatively, in particular dialog structures, each system can subscribe to the accessible resources and periodically, or when accessing, update the skeleton meta-information. Furthermore, meta-information servers can perform the spidering and provide the skeleton information along with the meta-information. Coffman at 41:29-34 During the registration phase, they will exchange upon request list of supported context. During the connection, these contexts are updated. Depending on the connection, the context is centralized or distributed across the devices (i.e., the network is negotiated). Coffman at 20:22-26

Claim 27 of the '468 Patent	Exemplary Citations to Coffman
	The active master dialog manager (illustrated as dialog manger 219 in FIG. 9) will be responsible for managing the flow of I/O to the different managers to decide the active dialog and appropriately execute a query and update the context and history. Coffman at 16:29-33: The global history stores, for example, the information that is associated with all the applications and actions taking during a conversational session (i.e., the history of the dialog between user and machine for a current session (or from when the machine was activated). Coffman at 15:50-16:17: The task dispatcher/manager 221 dispatches and coordinates different tasks and processes that are spawned (by the user and machine) on local and networked conventional and conversational resources (explained in further detail below). The meta information manager 222 manages the meta-information associated with the system via a meta-information repository 218. The meta information manager 218 and repository 218 collect all the information typically assumed known in a conversational interaction but not available at the level of the current conversation. Examples are: a-priori knowledge: cultural, educational assumptions and persistent information: past request, references, information about the user, the application, news, etc. It is typically the information that needs to be preserved and persist beyond the length/life of the conversational history/context and the information that is expected to be common knowledge for the conversation and therefore, has never been defined during the current and possible past conversational interactions. Also, as described below, shortcuts to commands, resources and macros, etc. are managed by the meta-information manager 222 and stored in the meta information repository 218. In addition, the meta-information repository 21 includes a user-usage log based on user identity. It is to be appreciated that services such as conversational help and assistance, as well as some dialog prompts (introduction, questions, feedback etc) provided by the CVM system can be tailored based on the usage history of the user as stored in the meta-information repository 218 and associated with the application. If a user has been previously interacting with a given application, an explanation can be reduced assuming that it is familiar to the user. Similarly, if a user commits many errors, the

Claim 27 of the '468 Patent	Exemplary Citations to Coffman
	explanations can be more complex, as multiple errors is interpreted as user uncertainty, unfamiliarity, or incomprehension/misunderstanding of the application or function. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

5. Claim 28

Claim 28 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 27, wherein the knowledge-enhanced speech recognition engine determines a most likely context for the follow-up multi-modal input using the updated context stack.	Coffman discloses "wherein the knowledge-enhanced speech recognition engine determines a most likely context for the follow-up multi-modal input using the updated context stack[.]" As one example, the system of Coffman that performs one or more of the method steps of claim 19 can be the knowledge-enhanced speech recognition engine. *E.g.*, Coffman at 31:23-30: Completion/search for the active context is performed from context to context down the stack. That is, new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user. As soon a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active.

Claim 28 of the '468 Patent	Exemplary Citations to Coffman
	See also citations of claims 26 and 27. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

6. Claim 29

Claim 29 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 27, wherein the speech recognition engine creates a speech-based transcription of the follow-up natural language utterance using the updated semantic knowledge-based model.	Coffman discloses "wherein the speech recognition engine creates a speech-based transcription of the follow-up natural language utterance using the updated semantic knowledge-based model[.]" *E.g.*, Coffman at 4:47-64: In another aspect, the CVM provides the capability to have natural dialog with NLU, NLG, contexts and mixed-initiatives sorted across multiple tasks, processes and discourses (with multiple domains). A conversational input interface is provided whereby a set of multi-mode input streams are each transcribed into an ASCII command or query (i.e., lists of attribute-value pairs or n-uples). Each input entity (command, NLU query field or argument unit (isolated letter, word, etc.) is associated with time-marks and appended accordingly to a compounded input stream. Two or more stream having the same time-marks are prioritized based on when each input stream contributed previously or the priority that each application/input stream received on the basis of

Claim 29 of the '468 Patent	Exemplary Citations to Coffman
	the context history. Compounded inputs are checked against possible FSG and dictionaries and optionally fed back to the user. Each resource exchanges their conversational capabilities and the input stream is tailored to only exchange relevant information. Coffman at 28:39-57: In accordance with the present invention, the input problem is treated as a merge of the output of multiple decoders, ASCII transcription or a list of attribute value n-uples. Each input stream is converted into its ASCII transcription and aligned with input time-marks via speech recognition processing 604. When different input stream are associated to the same task, the transcripts are merged as follows. First, commands and queries are sorted based on the time marks and appended to a single data stream. Command formulation can be checked against FSG rules and re-sorted to satisfy the grammar rules. NLU queries do not necessarily require re-sorting. For NLU queries, the symbolic fields are filled for each stream, then compounded at the level of the final input stream. Arguments such as spelling and alpha-numeric code do not exploit grammar rules or NLU to solve ordering ambiguity. Time-marks are used similarly to build a unique stream. However, the input is fed back to the user for confirmation, with possible pre-filtering using a dictionary or FSG rule book which is application-dependent. Coffman at 30:36-50: The task dispatcher/controller 402 divides each command/process into multiple actions, starts the associated threads/processes with the appropriate priority and relates/inserts them within the context stack 405. The task dispatcher 402 allocates each resource and shares them between the different spawned actions, and controls handles and streams to and from the resources. Based on the modality (pointer, keyboard, file, speech), the task dispatcher 402 redirects the stream to the appropriate conversational subsystems or conventional subsystem with speech inputs being transcribed/understood. The output of these subsystems is run down the context stack 405 to extract the active query and complete it. On the other hand, outputs are queued based on the priority levels of each task and dispatched sequentially to the output resource.

Claim 29 of the '468 Patent	Exemplary Citations to Coffman
	See also citations of claim 27. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

7. Claim 30

Claim 30 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 19, wherein the identified domain agent processes the request by querying one or more local or network information sources.	Coffman discloses "wherein the identified domain agent processes the request by querying one or more local or network information sources[.]" *E.g.*, Coffman at 6:56-67: In yet another aspect, conversational help, manuals and support is provided through a ubiquitous coordinated conversational interface, using local and remote resources, usage history of a user and agents to complete request, guide through procedure, search for information and upgrade/install new applications. In addition, help information can be accessed using NLU queries to access the help information or on the basis of the meta-information associated to the current user (history) and on the basis of the arguments that are missing or modified using mixed initiative. The dialog provided by each application is tuned to the preferences or level of expertise of the user. Coffman at 42:1-32:

Claim 30 of the '468 Patent	Exemplary Citations to Coffman
	When multiple devices are conversationally connected, they will coordinate their interfaces so that all the devices can be controlled through the universal CUI. This concept may be illustrated by the following example. Assume that you are driving home one night and remember that your spouse asked you to buy some goods at a new grocery store. After finding the message on your answering machine, you rapidly transcribed it into a memo on your desktop using a speech recognition software. However, you forgot to print it or transfer it on your PDA. It does not matter if your desktop PC runs a CVM since you have, in your car, a conversational PDA, a conversational car PC (PVA, Personal Vehicle Assistant) and a conversational smart phone. Further assume that the PVA runs an embedded CVM while the two other applications are conversationally aware, i.e., you can control them through the CVM running on the PVA. You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. Coffman at 42:33-39: You now instruct the PVA to guide you turn by turn to the store. Your position is computed, the location of the store is fetched, locally or from a server, and an itinerary is computed to take into account the latest traffic information. At any time, you can request navigation information about where you are, what to do next, how far to go or even request a different itinerary.

Claim 30 of the '468 Patent	Exemplary Citations to Coffman
	To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.

8. Claim 32

Claim 32 of the '468 Patent	Exemplary Citations to Coffman
The method of claim 19, wherein the identified domain agent processes the request by directing a command to one or more local or remote devices.	Coffman discloses "wherein the identified domain agent processes the request by directing a command to one or more local or remote devices[.]" *E.g.*, Coffman at 42:1-32: When multiple devices are conversationally connected, they will coordinate their interfaces so that all the devices can be controlled through the universal CUI. This concept may be illustrated by the following example. Assume that you are driving home one night and remember that your spouse asked you to buy some goods at a new grocery store. After finding the message on your answering machine, you rapidly transcribed it into a memo on your desktop using a speech recognition software. However, you forgot to print it or transfer it on your PDA. It does not matter if your desktop PC runs a CVM since you have, in your car, a conversational PDA, a conversational car PC (PVA, Personal Vehicle Assistant) and a conversational smart phone. Further assume that the PVA runs an embedded CVM while the two other applications are conversationally aware, i.e., you can control them through the CVM running on the PVA.

Claim 32 of the '468 Patent	Exemplary Citations to Coffman
	You can instruct the PVA to dial in your PC using the phone. Once the connection is established, you are authenticated by voice and you find by voice the memo by simply requesting the "grocery list" that you had previously created, without having to remember the file name or the directory or browse through your directory to eventually select the appropriate file. You may need to confirm the selection if your PC CVM requests it. You can issue another query—"it should be synchronized with my PDA!—and the file is appropriately linked to be transferred to your PDA at the next synchronization. One last command—"Do it!"—and your PC gives up and lets the PVA handle that ambiguous query. The PVA understands your desire to synchronize the PDA and the PC based on your previous conversation. After possible confirmation, the synchronization is performed and the grocery list is stored on your PDA, ready for later use. Coffman at 42:33-39: You now instruct the PVA to guide you turn by turn to the store. Your position is computed, the location of the store is fetched, locally or from a server, and an itinerary is computed to take into account the latest traffic information. At any time, you can request navigation information about where you are, what to do next, how far to go or even request a different itinerary. To the extent it is found that Coffman does not expressly disclose this limitation, the limitation is inherent. And to the extent it is found that Coffman does not anticipate this claim, Coffman renders it obvious, standing alone, in combination with the knowledge of a person of ordinary skill in the art, and/or in combination with one or more other prior art references identified in the Cover Pleading or herein, for the reasons identified herein or in the Cover Pleading. A person of ordinary skill in the art would have been motivated to combine Coffman with any of the other prior art references identified in the Cover Pleading or herein with a reasonable expectation of success because each of these references contains familiar elements within the similar fields of endeavor to address similar known problems, according to known methods which would yield predictable results.