

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

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

v.

VB ASSETS, LLC,
Patent Owner

**IPR2025-00866
U.S. Patent No. 8,073,681**

**DECLARATION OF HOUMAN HOMAYOUN, PH.D. IN SUPPORT OF
PATENT OWNER'S PRELIMINARY RESPONSE**

TABLE OF CONTENTS

I.	INTRODUCTION	3
II.	QUALIFICATIONS.....	5
III.	LEVEL OF ORDINARY SKILL IN THE ART.....	12
IV.	CLAIM CONSTRUCTION.....	12
V.	PROSECUTION HISTORY	13
VI.	OVERVIEW OF OPINIONS.....	17
VII.	GROUND 1	18
A.	Overview of the Asserted References	18
1.	SmartKom (Ex. 1005).....	18
2.	Kobsa (Ex. 1006)	19
B.	The SmartKom + Kobsa Combination Fails to Describe a Coherent System that Includes all the Elements of the '681 Patent	20
C.	A POSITA Would Not be Motivated to Combine SmartKom and Kobsa to Reach the Claimed Invention.....	35
D.	The Petition Fails to Show that the SmartKom + Kobsa Combination Discloses or Teaches “short-term shared knowledge”	38
VIII.	FOUNDATIONS 2-4.....	41
A.	Overview of the Asserted References	42
1.	Barbara (Ex. 1007).....	42
2.	Ross (Ex. 1008).....	45
B.	A POSITA Would Not Be Motivated to Combine Barbara and Ross	46
C.	The Barbara + Ross Combinations Fails to Disclose or Teach “determining an intended meaning” because Barbara is based on human assistance	52

I, Houman Homayoun declare as follows:

I. INTRODUCTION

1. I have personal knowledge regarding the facts discussed herein, I am over the age of 18, and I am able to competently testify.

2. I understand that Samsung Electronics Co. Ltd. and Samsung Electronics America, Inc. (collectively, “Petitioners” or “Samsung”) challenged the patentability of U.S. Patent No. 8,073,681 (the “’681 Patent”) in a petition for *Inter Partes* Review (“IPR”).

3. I have been retained as an expert by VB Assets LLC (“Patent Owner”), to provide opinions regarding the patentability of U.S. Patent No. 8,073,681 (the “’681 Patent”).

4. My opinions regarding the patentability of the ’681 Patent are set forth herein. My opinions are based on the many documents I have reviewed, including the ’681 Patent, its file history, the petition (Paper No. 3, referred to as the “Petition”) and the exhibits and materials cited therein (Exs. 1001 – 1031), including the Declaration of Stuart Lipoff (Ex. 1003).

5. The opinions set forth herein are based on the knowledge and information currently available to me. I reserve the right supplement and/or amend the opinions set forth herein.

6. Further, I understand that Petitioners assert four invalidity Grounds in this IPR based six on references (i) SmartKom: Foundations of Multimodal Dialogue Systems (2006), (ii) User Models in Dialog Systems, Kobsa & Wahlster eds. (1989), (iii) U.S. Patent App. Publication No. 2004/0101198 to Barbara (“Barbara”), (iv) U.S. Patent App. Publication No. 2002/0173960 to Ross (“Ross”), (v) U.S. Patent App. Publication No. 2005/0278180 to O’Neill (“O’Neill”), and (vi) U.S. Patent App. Publication No. 2010/0125458 to Franco (“Franco”). The four invalidity grounds and the corresponding challenged claims are set forth in the table below:

Ground		Claims	Prior Art
1	103	1-9, 11-21, 23-33, 35-36	SmartKom+Kobsa
2	103	1, 5-8, 11-13, 17-20, 23-25, 29-32, 35-36	Barbara+Ross
3	103	2-4, 9, 14-16, 21, 26-28, 33	Barbara+Ross+O’Neill
4	103	10, 22, 34	Barbara+Ross+Franco

7. I understand that Patent Owner will submit a Preliminary Response to the Petition, and I have been asked review and analyze the Petition and accompanying exhibits and provide my opinions in response. I understand that Patent Owner’s Preliminary Response will be submitted prior to institution. To the extent the IPR is instituted, I reserve all rights to submit a declaration in support of Patent Owner’s response.

8. For the reasons set forth below, it is my opinion that the Petition fails to demonstrate that claims 1-36 are obvious in view of the asserted grounds.

9. I am being compensated at a rate of \$400 per hour for my study and other work in this matter. I am also being reimbursed for reasonable and customary expenses associated with my work in this investigation. My compensation is not contingent on the outcome of this matter or the specifics of my testimony.

II. QUALIFICATIONS

10. My curriculum vitae is attached as Exhibit 2026. A summary of my qualifications relevant to this proceeding is provided below.

11. I am a Professor of Electrical and Computer Engineering at the University of California, Davis. Prior to my current position, I was an Associate Professor at George Mason University (GMU). I am the director of GMU's Accelerated, Secure, and Energy-Efficient Computing Laboratory (ASEEC). Prior to joining GMU, I spent two years at the University of California, San Diego, as NSF Computing Innovation (CI) Fellow awarded by the CRA-CCC.

12. I hold a Ph.D. in Computer Science from the University of California, Irvine, which I earned in 2010, an M.S. in computer engineering from the University of Victoria, which I earned in 2005, and a B.S. in electrical engineering from the Sharif University of Technology, which I earned in 2003.

13. I serve as the director of the National Science Foundation Center for Hardware and Embedded Systems Security and Trust (CHEST).

14. I worked in NOVELICS startup company as a design architect from January 2007 to October 2008. I designed a reconfigurable and programmable processor referred to as Built-in Self-test (BIST) to test various memory architecture.

15. My research focuses on applied machine learning and AI, large language models, hardware security and trust, data-intensive computing, and heterogeneous computing.

16. I have directed over \$20M in funded research projects. My team at UC Davis and prior to that at George Mason University, has successfully completed more than 20 projects funded by DARPA, NIST, NSF and, AFLR, DHS, Intel, General Motors, and a consortium on more than 20 companies which are member of CHEST center. Several of these projects are related to using large language models and or designing large language models

17. From 2017 to 2021 I served as an Associate Editor of IEEE Transactions on VLSI. I served as TPC Co-Chair for GLSVLSI 2018. I also served as the general chair of GLSVLSI 2019.

18. I am the co-author of the textbook Machine Learning for Computer Scientists and Data Analysts, which covers recommender systems, inference

modeling, deep learning architectures, and deployment strategies for AI across software and hardware platforms.

19. My current research focuses on applied machine learning and AI, including deep learning, neural networks, natural language processing, user-aware computing, and conversational AI. I have published more than 100 peer-reviewed papers in these areas.

20. My research contributions include: "A Fast Method to Fine-Tune Neural Networks for the Least Energy Consumption on FPGAs," enabling efficient on-device AI for edge-based speech and audio processing, "Energy-Efficient and Adversarially Robust Machine Learning with Selective Dynamic Band Filtering," improving robustness in processing sensor signals, including voice, "Addressing Stereotypes in Large Language Models: A Critical Examination and Mitigation Approach," relevant to fairness, personalization, and user modeling in conversational AI, "A Hardware Accelerator for Language-Guided Reinforcement Learning," bridging natural language instructions and control logic for interactive agents, "Automatic Detection of Respiratory Symptoms Using a Low Power Multi-Input CNN Processor," advancing speech-driven health applications.

21. I teach both undergraduate and graduate courses at UC Davis. My Embedded Systems course includes projects on speech recognition applications for microcontrollers, integrating real-time inference, user interaction, and HCI design.

My Applied Machine Learning course covers NLP, transformers, LLMs, and conversational AI with emphasis on practical implementation and user modeling.

III. LEGAL PRINCIPLES

22. For my opinions in this declaration, I understand that it requires applying various legal principles. As I am not an attorney, I have been informed and understand from counsel about various legal principles that involve my analysis. I have used my understanding of those principles in forming my opinions. I summarize those principles as I understand them below.

23. For example, I understand that Petitioners bear the burden of proving unpatentability in this proceeding by a preponderance of the evidence. I understand that this preponderance of the evidence standard means that Petitioners must show that unpatentability is more probable than not

24. I understand that a claimed invention is not patentable under 35 U.S.C. § 103 if it is obvious. A patent claim is unpatentable if the claimed invention would have been obvious to a person of ordinary skill in the field (which I discuss below) at the time the claimed invention was made. This means that even if all of the requirements of the claim cannot be found in a single prior art reference that would anticipate the claim, a person of ordinary skill in the relevant field who knew about all this prior art would have come up with the claimed invention.

25. I have been informed and understand that the obviousness analysis requires evaluating the claim language as a whole, including comparing each and every limitation of the properly construed claim language to the prior art.

26. I further understand that the ultimate conclusion of whether a claim is obvious should be based upon several factual determinations. That is, a determination of obviousness requires inquiries into: (1) the level of ordinary skill in the field; (2) the scope and content of the prior art; (3) what difference, if any, existed between the claimed invention and the prior art; and (4) any objective factors bearing on obviousness.

27. I understand that, in determining the level of ordinary skill in the field that someone would have had at the time the claimed invention was made, I should consider: (1) the levels of education and experience of persons working in the field; (2) the types of problems encountered in the field; and (3) the sophistication of the technology.

28. I understand that, in determining the scope and content of the prior art, in order to be considered as prior art, a reference must be analogous art. I also understand that to be analogous art, the reference must be either (1) from the same field of endeavor as the claimed subject matter, regardless of the problem addressed, or (2) if not in the same field of endeavor, reasonably pertinent to the particular problem with which the inventor is involved. I am also familiar with the premise that

for a reference to be reasonably pertinent, it must have logically commended itself to an inventor's attention at the time of invention.

29. I understand that a patent claim composed of several elements is not proved obvious merely by demonstrating that each of its elements was independently known in the prior art. In evaluating whether such a claim would have been obvious, I may consider whether there is a reason that would have prompted a person of ordinary skill in the field to combine the elements or concepts from the prior art in the same way as in the claimed invention.

30. I understand that in order to show obviousness based on a single reference or a combination of references, a particular motivation to modify the reference or combine the teachings in the references, and a reasonable expectation of success must be shown.

31. I understand that the art must evidence a motivation to combine or modify the independently known elements to arrive at the claimed invention and an explanation as to how or why the references would be combined to arrive at the claimed invention to solve the particular problem. I understand that a challenger must show that there was a reason that would have prompted a person of ordinary skill in the relevant field to combine the elements in the way the claimed new invention does in an obviousness determination. I also understand that it is not

enough to prove obviousness merely by a showing that the references were capable of combination without a suggestion of the desirability of the modification.

32. I further understand that there is no single way to define the line between true inventiveness on the one hand (which is patentable) and the application of common sense and ordinary skill to solve a problem on the other hand (which is not patentable). For example, market forces or other design incentives may be what produced a change, rather than true inventiveness. I may consider whether the change was merely the predictable result of using prior art elements according to their known functions, or whether it was the result of true inventiveness. I may also consider whether there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent. I may consider whether the innovation applies a known technique that had been used to improve a similar device or method in a similar way. I may also consider whether the claimed invention would have been obvious to try, meaning that the claimed innovation was one of a relatively small number of possible approaches to the problem with a reasonable expectation of success by those skilled in the art.

33. I also understand, however, that I must be careful not to determine obviousness using the benefit of hindsight; many true inventions might seem obvious after the fact. I should put myself in the position of a person of ordinary skill

in the field at the time the claimed invention was made and I should not consider what is known today or what is learned from the teaching of the patent.

34. Finally, I understand that any obviousness rationale for modifying or combining prior art must include a showing that a person of ordinary skill would have had a reasonable expectation of success. I further understand that whether a proposed modification or combination of the prior art has a reasonable expectation of success is determined at the time the invention was made.

III. LEVEL OF ORDINARY SKILL IN THE ART

35. I understand that the Petition asserts that a person of ordinary skill in the art for the '681 Patent ("POSITA") would have the following level of knowledge:

A bachelor's degree in computer science, computer engineering, electrical engineering, or related field in computing technology, and two years of experience with automatic speech recognition and natural language understanding, or equivalent education, research experience, or knowledge.

36. I understand that the Board adopted this definition for the relevant level of ordinary skill in the art in prior IPR2020-01367. For purposes of this declaration, I do not offer an independent definition of a POSITA, and have applied the definition set forth above. I reserve the right to offer additional opinions regarding the level of ordinary skill in the art should this IPR be instituted.

IV. CLAIM CONSTRUCTION

37. Based on the representations in the Petition, I understand that the term "speech recognition engine," which appears in claims 7, 10, 19, 22, 31, and 24, was

previously construed as “software or hardware that recognizes the words or phrases in the natural language utterance” based on agreement of the parties in a prior litigation. I have considered and applied this construction in my analysis and opinions set forth herein. I have otherwise applied the plain meaning of the challenged claims. At this time, I do not offer any additional opinions regarding the proper construction of any of the claims, but I reserve the right to further address the proper construction of the claims if IPR is instituted.

V. PROSECUTION HISTORY

38. As set forth above, I have reviewed the file history of the '681 Patent, and below I provide an overview of the file history, including aspects of the file history that are relevant to my opinions set forth herein.

39. The application that resulted in the issuance of the '681 Patent was filed in 2006, and the application was published in 2008.

40. The examiner issued a non-final rejection in June 2010, rejecting certain of the pending claims, among other statutory grounds, under § 103 over Kargman, Morin, and in further view of Bielbly. Ex. 1002, 175-187. The examiner stated that [c]laims 12, 26, 40 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Kargman, as applied to Claims 1, 15, and 29, above, and further in view of Bielby et al. (US 5,488,652), hereinafter Bielby.” Ex. 1002, 185. Examiner stated that “Bielby suggests learning that the determined intended

meaning is incorrect . . . further updating the shared knowledge to include knowledge that the determined meaning is incorrect . . . and determining a second intended meaning for the at least one utterance based on the further updated shared knowledge[.]” Ex. 1002, 185-186.

41. In response to the rejection, the applicant amended the claims and submitted remarks with the amendment. Ex. 1002, 212-237. In the remarks, Applicant argued “Kargman fails to disclose, teach, or suggest at least the combination of features that include “receiving an utterance at a voice input device, wherein the voice input device receives the utterance in a current conversation with a user” and “determining an intended meaning for the utterance,” wherein a conversational speech engine determines the intended meaning for the utterance from short-term shared knowledge about the current conversation and long-term shared knowledge about the user, including one or more past conversations with the user, as recited in amended independent claim 1, for example.” Ex. 1002, 232. Applicant also argued that “Bielbly further fails to cure the foregoing deficiencies of Kargman for at least the reason that although Bielbly generally describes a system that can ask a caller about whether a call entity recognized an utterance is incorrect (col. 18, lines 1-30), Bielbly does not indicate that ‘training speech recognition algorithms’ in the system described therein results in the system using trained speech recognition algorithms to determine the intended meaning for the utterance.” Ex.

1002, 235. Applicant provided that, “rather than attempting to independently disambiguate or otherwise determine the intended meaning for the utterance, the system in Bielbly prompts the user to confirm or disconfirm whether certain words were correctly recognized.” Ex. 1002, 235. Thus, Applicant argued that “because the system described in Bielbly relies upon user feedback to determine whether an utterance was correctly recognized, Bielbly does not disclose, teach, or suggest that the *system* determined an ‘intended meaning’ for an utterance.” Ex. 1002, 235 (emphasis added).

42. The Examiner issued a final rejection in November 2010. Ex. 1002, 249-277. However, the Examiner’s response to Applicant’s arguments in this rejection was directed to limitations other than “determining the intended meaning.” Ex. 1002, 253-255. In its response to Applicant’s arguments, the Examiner stated that “Applicant’s remaining arguments are directed to the secondary references not teaching ‘determining the intended meaning’. Since the secondary references are not applied to teach this limitation, the examiner maintains the previous prior art rejections under the secondary references, directed to what the secondary references were previously applied to teach.” Ex. 1002, 256. Thus, as before, the Examiner stated that “Bielby suggests wherein determining the intended meaning for the utterance includes generating multiple possible interpretations for the utterance, wherein an initial interpretation for the utterance comprises one of the multiple

possible interpretations having a highest degree of certainty . . . and updating the short-term shared knowledge about the current conversation to remove the initial interpretation from the multiple possible interpretation in response to determining that the initial interpretation was incorrect, wherein the intended meaning determined from the short-term shared knowledge and the long-term shared knowledge comprises one of the multiple possible interpretations having a next highest degree of certainty” and further that “Bielby contained a comparable user-machine interaction device where multiple interpretations of input speech can be made and has been improved in the same way as the claimed invention by incorporating a process to accommodate recognition-errors.” Ex. 1002, 272-273.

43. The Applicant responded by amending the claims in January 2011 and submitting remarks with the amendment. Ex. 1002, 299-321. In its remarks, Applicant argued that Kargman “fails to disclose, teach or suggest . . . ‘establishing [an] intended meaning [for the utterance] within the identified context, wherein [a] conversational speech engine establishes the intended meaning within the identified context to disambiguate an intent that the user had in speaking the one or more words that have the different meanings in different contexts,’ as recited in claim 1[.]” Ex. 1002, 318. Specifically, applicant again argued that “Bielbly further fails to cure the foregoing deficiencies of Kargman for at least the reason that although Bielbly generally describes a system that can ask a caller about whether a call entity

recognized in an utterance is incorrect (col. 18, lines 1-30), Bielbly does not describe using results from the ‘speech recognition algorithms’ trained in the system described therein to determine or otherwise disambiguate an intended meaning or intent associated with certain words that have different meanings in different contexts.” Ex. 1002, 319. Thus, Applicant argued that “Bielbly relates, at best, to determining what words or words were actually said, . . . but does not disclose, teach, or suggest establishing or otherwise disambiguating the meaning associated with certain words that were presumably recognized correctly based on the context in which the words were said.” Ex. 1002, 319.

44. The examiner issued a Notice of Allowance on August 15, 2011. Ex. 1002, 919, and an issue notification was filed on November 16, 2011. Ex. 1002, 1012.

VI. OVERVIEW OF OPINIONS

45. It is my opinion that the Petition’s asserted combination of SmartKom and Kobsa does not render obvious claims 1-9, 11-21, 23-33, and 35-36.

46. It is my opinion that the Petition’s asserted combination of Barbara and Ross does not render obvious claims 1, 5-8, 11-13, 17-20, 23-25, 29-32, and 35-36.

47. It is my opinion that the Petition’s asserted combination of Barbara, Ross, and O’Neill does not render obvious claims 2-4, 9, 14-16, 21, 26-28, and 33.

48. It is my opinion that the Petition’s asserted combination of Barbara, Ross, and Franco does not render obvious claims 10, 22, and 34.

VII. GROUND 1

49. The first ground of the Petition asserts that the Challenged Claims are obvious over SmartKom (Ex. 1005) and Kobsa (Ex. 1004). I disagree that Ground 1 of the Petition demonstrates that the Challenged Claims are obvious over SmartKom and Kobsa.

A. Overview of the Asserted References

1. SmartKom (Ex. 1005)

50. SmartKom is a textbook comprising 38 different subparts, which are authored by different contributors from a list of over 70. SmartKom describes the “present the theoretical and practical foundations of multimodal dialogue systems using the results of [the] large-scale project SmartKom as the background for [the] discussion.” Ex. 1005, IX – XIV, 5. SmartKom provides a “comprehensive overview of the broad spectrum of results of the research conducted in SmartKom.” Ex. 1005, VI. SmartKom was a large-scale project funded by the German Government designed to explore multimodal dialogue systems. The project spanned four years and the contributors and researchers included a consortium of multiple industrial companies, research institutes and universities. Thus, the sub-parts of SmartKom are directed to theories and principles in multimodal dialogue systems, and describes

corresponding aspects of the so-called SmartKom system or architecture that was developed by dozens of different researchers at different times.

51. The Petition relies extensively on SmartKom’s intention recognition disclosure, which is shown in the Figure below.

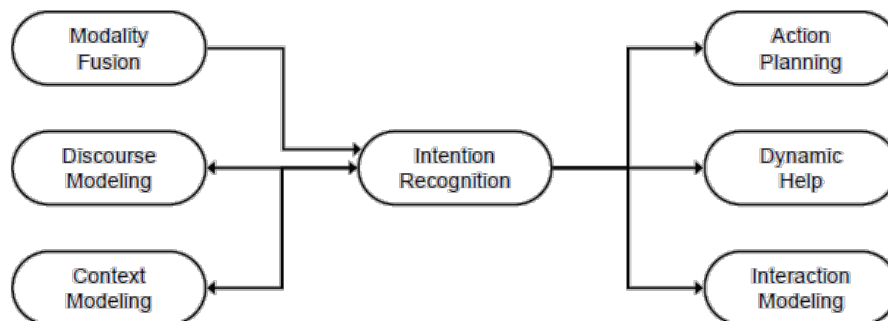


Fig. 1. Section of the multimodal interaction from the intention recognition point of view

Ex. 1005, 286.

52.

2. Kobsa (Ex. 1006)

53. Kobsa (Ex. 1006) is also a textbook directed to documenting the results of an international workshop on user modeling that brought together 25 researchers. The stated intention of Kobsa is to form a “collection of unintegrated conference papers” and to “provide a rather coherent survey of the field of user modeling.” Ex. 1006, V. Kobsa includes several Parts, each comprising numerous subparts, including a survey of user modeling, building user models, exploiting user models, and shortcomings of current models, prospects for the future. The Petition relies on

broad ranging disclosure of the various subparts of Kobsa. For example, as part of Kobsa's survey of the field of user modeling, scattered throughout are descriptions of different systems and models directed to different implementations. For example, Kobsa discloses a system called GRUNDY as an example of a system that uses an initial individual user model. GRUNDY is a system that recommends novels to people to read. Kobsa also discloses difference general user modeling systems, including the GUMS system, the TRUMP system, and the BGP-MS system.

**B. The SmartKom + Kobsa Combination Fails to Describe a
Coherent System that Includes all the Elements of the '681 Patent**

54. The first ground of the Petition asserts that the Challenged Claims are obvious over SmartKom (Ex. 1005) and Kobsa (Ex. 1004). I have reviewed SmartKom and Kobsa, and each of these references are broad-ranging and reflect the work of many different researchers over time. In my opinion, Petitioners rely on disparate disclosures in SmartKom and Kobsa such that the proposed combination does not disclose or render obvious a coherent system that satisfies each element of the '681 Patent claims.

55. SmartKom, on the one hand lists over 70 different contributors who, in 38 different sub-parts, “present the theoretical and practical foundations of multimodal dialogue systems using the results of [the] large-scale project SmartKom as the background for [the] discussion.” Ex. 1005, IX – XIV, 5. The SmartKom reference is self-described as providing a “comprehensive overview of the broad

spectrum of results of the research conducted in SmartKom.” Ex. 1005, VI. The SmartKom project was developed over a four-year period, and the final SmartKom demonstrator was presented in June 2003. Ex. 1005, 35. The plan for the project was defined by nine subprojects assigned to twelve different research groups. Ex. 1005, 31. This project was coordinated by the German Research Center for Artificial Intelligence, and comprised a consortium of six companies, one research institute, and three universities. Ex. 1005, V. As SmartKom itself describes, there is disclosure of “theoretical and practical foundations of multimodal dialogue systems,” as well as disclosure of different components of concepts for implementation based on the SmartKom project developed over time, and corresponding examples of different possible implementations. This is important because, in some instances, Petitioners rely on generalized disclosures not a part of SmartKom’s description of its implementation, or other, unasserted references, rather than the description of the implementation of the SmartKom system or architecture. Moreover, the Petition relies on different aspects of SmartKom developed by different people that are described for different implementations. The Petition fails to tie these different aspects together to show that they form a system in accordance with the challenged claims.

56. While SmartKom describes the implementation of different aspects of a system or architecture for different applications, Kobsa, on the other hand,

describes itself as a “collection of unintegrated conference papers” from a conference of 25 researchers intended to provide a survey of the field of user modeling. Ex. 1006, V-VI. Thus, Kobsa is meant to broadly describe the aims and the scope of the field of user modeling and describes different systems that have been implemented and their results to achieve a broad description of the field. Kobsa describes the scope of the field of user modeling by reference to many different systems that have been implemented.

57. As set forth in several examples below, in my opinion, the Petition relies on disparate disclosure and fails to identify a coherent system in accordance with the claimed invention.

58. First, because of the broad-ranging nature of SmartKom and Kobsa, each of which was developed over time by different people, neither SmartKom nor Kobsa describe a single coherent system. The Petition cherry picks pieces from throughout each reference to satisfy different claim limitations, but fails to sufficiently explain how those cherry-picked disclosures would come together in a coherent system or method that includes all the elements of the challenged claims. The challenged claims recite “a cooperative conversational voice user interface” comprising “a voice user input device,” configured in a certain way, and “a conversational speech engine with one or more processors” configured a certain way.. This combination of limitations in the challenged claims are not arbitrary or

ad hoc, but instead represent a non-obvious integration of components of methods to address a specific problem.

59. The Petition, however, attempts to rebuild the claimed invention by cherry-picking unrelated pieces from SmartKom and Kobsa using the challenged claims as a guide. Using this hindsight-driven reasoning is not how design of dialogue systems works, and it is not how a POSITA would approach the problems identified in the '681 Patent at the time. Specifically, the Petition extracts individual features from separate prior art systems, or generalized concepts applicable to all prior art systems and retrofits them into a patchwork of the '681 Patent without acknowledging the vast combinatorial design space inherent in dialog system architecture.

60. To illustrate this technically, Claim 25 of the '681 Patent recites a system for providing a cooperative conversational voice user interface, which comprises a voice input device configured in a specific way, and a conversational speech engine which includes “one or more processors” configured to: (i) “accumulate short-term shared knowledge about the current conversation, wherein the short-term shared knowledge includes knowledge about the utterance received during the current conversation;” (ii) “accumulate long-term shared knowledge about the user, wherein the long-term shared knowledge includes knowledge about one or more past conversations with the user;” (iii) “identify a context associated

with the utterance from the short-term shared knowledge and the long-term shared knowledge;” (iv) “establish an intended meaning for the utterance within the identified context to disambiguate an intent that the user had in speaking the one or more words that have the different meanings in the different contexts; and” (v) “generate a grammatically or syntactically adapted response to the utterance based on the intended meaning established within the identified context.” Ex. 1001, cl. 25.

61. A POSITA designing a system providing a voice user interface would understand that each of these limitations could correspond to modules that could be configured in at least ten different ways (via alternative algorithmic approaches, interface protocols, or processing assumptions). The resulting potential design permutations exceed 1,000,000. This design complexity underscores the improbability of arriving at the exact limitations of the ’681 Patent while reviewing references such as SmartKom and Kobsa with far-reaching disclosures detailing many different concepts and systems absent the hindsight of the claims as a guide. In my opinion, even if the Petition identifies the individual building blocks as existing in the literature in either SmartKom or Kobsa, the Petition fails to show how those building blocks are organized, constrained, and functionally tied to one another to reach the claimed invention recited in the ’681 Patent.

62. For example, the Petition relies on SmartKom’s disclosure of an intention recognizer for some aspects of the claim limitations, and SmartKom’s

disclosure of action planning for other claim limitations. Petition, 23 (“In the SmartKom architecture, the intention recognizer establishes the intended meaning of an utterance within the identified context—i.e., disambiguates an intent that the user had in speaking the one or more words that have different meanings in different contexts.”); at 39 (“The action planner ‘*generates a response to the utterance*’ through development of an action plan ‘specify[ing] a possible course of subsequent communicative acts to reach’ a goal articulated by the user in the utterance.”).

63. These are examples of different aspects of SmartKom, that, while funded under the same research umbrella, were developed and described by different researchers, and described using different implementations. For example, the Petition states that SmartKom’s “intention recognizer establishes the intended meaning of an utterance within the identified context—i.e., disambiguates an intent that the user had in speaking the one or more words that have different meanings in different contexts.” Petition, 36. The functionality of the intention recognizer is described in the Petition using the example from SmartKom where a user request “Is there a movie with Nicole Kidman?” leads to two intention hypotheses: one where the user requests information about the current TV program, and one where the user requests information about the current cinema program. Petition, 28-32 (citing Ex. 1005, 263-265). This description appears in the chapter of SmartKom that presents “a formal description of overlay—the fundamental operation for discourse

processing in SmartKom.” Ex. 1005, 265. This chapter was authored by Jan Alexandersson, Tilman Becker, and Norbert Pfleger. Ex. 1005, 255. On the other hand, the Petition relies on SmartKom’s description of the action planner as disclosing “generat[es the] response to the utterance.” Petition, 39-40. The functionality of the action planner is described in the Petition using entirely different implementations from the Nicole Kidman implementation for determining whether the utterance requests information about a TV program or cinema program in the intention recognizer. Namely, the Petition relies on an action plan for sending a fax and an action plan for an utterance to provide TV listings to a user where the response is based on whether a TV program broadcast was found or not found. Petition, 41-42. These examples of the action planning appear in the chapter of SmartKom titled “Plan-Based Dialogue Management for Multiple Cooperating Applications, which was authored by Markus Löckelt. Ex. 1005, 301.

64. In my opinion, the fact that the different individual components or modules disclosed in SmartKom that Petitioners rely upon for different claim limitations are authored by different researchers and directed to different implementations demonstrates that SmartKom does not disclose a single, coherent system, which serves as the backdrop for my opinions that the Petition fails to identify a coherent system that includes all the limitations of the asserted claims.

Specifically, as described above, SmartKom describes generalized principles of dialogue systems as well as the SmartKom the implementation and application for discrete components. Kobsa describes a survey of the field of user modeling, which includes the description of dozens of different, independent systems that necessarily have different components and different implementations. A POSITA would not be motivated to carefully select the different pieces of SmartKom and Kobsa that the Petition selects and expect to reach the claimed invention because there would be endless possibilities for implementation when selecting individual components for the resulting system.

65. Thus, in my opinion, the Petition fails to sufficiently explain how the cherry-picked disclosures of SmartKom and Kobsa would come together in a coherent system or method that includes all the elements of the challenged claims.

66. The first example of this is the Petition's reliance on both the disclosure of SmartKom that is focused on generalized descriptions of all dialogue systems, as well as its disclosure of SmartKom's implementation system. For example, the claims require "accumulating long-term knowledge about the user, wherein the long-term shared knowledge includes knowledge about one or more past conversations with the user," and "identifying . . . a context associated with the utterance from the short-term shared knowledge and the long-term shared knowledge." *See, e.g.*, cl. 1. For the limitations directed to "long-term shared knowledge", the Petition identifies

SmartKom’s disclosure of a user model that provides (i) “properties of the interlocutors”. Petition, 22. SmartKom’s teaching of the “properties of the interlocutors” as the content stored in the user model does not refer to the SmartKom system. Instead, this disclosure is a generalized “broad categorization of the types of context relevant to spoken dialogue systems” and their “content and respective knowledge stores” applicable to all systems. Ex. 1005, 274. SmartKom discloses that *in general* in human-to-human interactions, interlocutionary context is important in the process of employing “context-dependent resources” by the addressee to extract the meaning from the utterance that the speaker wanted to convey. This disclosure therefore does not refer to the SmartKom system that the Petition relies upon to satisfy other claim limitations. Instead, the disclosure simply refers to generalized foundations for all types of speech processing, including human-to-human conversations.

67. For other limitations the Petition relies on different aspects of what is described as the “SmartKom system” or “SmartKom architecture” described in various of the 38 different sub-parts of SmartKom. For example, the independent claims recite accumulating short-term shared knowledge and accumulating long-term shared knowledge, and identifying a context for the utterance *from* the short-term shared knowledge and long-term shared knowledge. Similarly, the independent claims require determining an intended meaning for the utterance, and

“grammatically or syntactically” adapting a response “*based on* the intended meaning within the identified context.” Where Petitioners rely on broad, generalized principles as I described above, such as the “properties of interlocutors,” rather than implementation aspects of the SmartKom system for “accumulating long-term knowledge,” but then subsequently rely on specific aspects or components of the SmartKom system for the claimed “long-term shared knowledge” in the limitations that require “identify[ing] a context” “based on the long-term shared knowledge” limitations, the Petition fails to identify a coherent system in accordance with the claims.

68. Furthermore, the Petition identifies three different possible disclosures in the SmartKom-Kobsa combination to disclose or teach “long-term knowledge”: (i) the generalized disclosure of a domain model in SmartKom; (ii) the generalized disclosure of a user model in SmartKom; and (iii) Kobsa’s disclosure of “long-term” user models that “describe relatively static characteristics of users.” Petition, 19-20. The Petition subsequently argues that SmartKom’s disclosure of an “interaction model” in the SmartKom system is a technique that meets the “accumulating long-term shared knowledge” limitation. Petition, 23. The interaction model in SmartKom operates in the SmartKom system to “compute information on the user” to be used for example, “to adapt the interaction style of the system to the interaction style of the user.” Ex. 1005, 287. However, as set forth above, the claims also require that

the claimed “long-term shared knowledge” is used to “identify a context.” Ex. 1001, cl. 1. The Petitioners fail to cite to any disclosure in SmartKom indicating that information computed on the user to adapt the interaction style (as performed by the interaction model) is used to identify context in accordance with the claims. Notably, Petitioners subsequently assert that the “identify[ing]” “a context . . . from the long-term shared knowledge is met by the intention recognizer process in the SmartKom system that engages the discourse modeler and the context modeler to enrich information in hypotheses with context knowledge.” Petition at 27-28. Petitioners state that “a hypothesis is enriched with “long-term shared knowledge” from the user model such as the user’s preferences or expertise in the identified context domain.” Petition, 33-34. Petitioners cite page 274 of SmartKom in support of this statement, but based on my review of this page, the only discussion of user models on this page is the generalized user model “broad categorization of the types of context relevant to spoken dialogue systems, their content and respective knowledge stores.” Ex. 1005, 274. There is no disclosure that teaches or suggests that user preferences or expertise are used as part of the context domain to identify a context associated with an utterance that includes one or more words having different meanings in different contexts. Accordingly, in my opinion, the Petition cherry picks these disclosures to satisfy different limitations, but fails to explain how each of these identified

disclosures relate to one another such that the asserted combination discloses or teaches a coherent system that includes all the elements of the challenged claims.

69. Furthermore, in my opinion, the combination of SmartKom and Kobsa fails to identify a coherent system in accordance with the claimed invention because Petitioners rely on different, and in many instances disparate, disclosures to satisfy the same claim limitations, without explaining how the combination of both disclosures would meet the claims.

70. For example, Petitioners allege that both the “interaction model” disclosed in SmartKom *and* the “General User Modeling System,” which can be “persistent,” in Kobsa satisfy the limitations requiring “accumulating long-term shared knowledge.” Petition, at 23. The Petition additionally identifies Kobsa’s disclosure of “implicit acquisition” of knowledge about a user by “eavesdropping on the user-system interaction” for these same limitations.

71. On the one hand, the interaction model in SmartKom operates in the SmartKom system to “compute information on the user” to be used for example, “to adapt the interaction style of the system to the interaction style of the user.” Ex. 1005, 287. On the other hand, Kobsa’s General User Modeling System (“GUMS”) for building long term models of individual users is aimed, not at interaction preferences of the user, but at representing the “knowledge and beliefs of individuals.” Ex. 1006, 417. Kobsa’s GUMS is acquired by allowing the application

for which it is used to select an initial stereotype for the user, which is another significant difference from SmartKom's interaction modeling. Ex. 1006, 418. Critically, Kobsa's GUMS system organization is such that the "user modeling system does not have access to the interaction between the application and the user." Ex. 1006, 418. The Petition fails to explain this inconsistency between the interaction modeling identified from SmartKom and the GUMS identified in Kobsa. Even further, Kobsa discloses four different possible aspects of eavesdropping or implicit acquisition: (i) presuppositions, (ii) misconception recognition, (iii) stereotype selection, (iv) use of default rules based on four different separate references. Ex. 1006, 416. The Petition fails, however, to address these four possible aspects, let alone explain how any one of these aspects of implicit acquisition could be implemented with SmartKom's interaction modeling. Accordingly, it is my opinion that none of these disclosures identified as satisfying the same challenged claim limitations fit together as describing the same concept or functionality, nor does the Petition explain how, for example, the General User Modeling System and/or the four different possible aspects of implicit knowledge acquisition would be incorporated with the interaction modeling in the SmartKom system. This is therefore an example where Petitioners have identified disparate disclosures from SmartKom and Kobsa, and failed to explain how the combination of these disclosures would satisfy the challenged claims.

72. Another example of the failure of Ground 1 to identify a coherent system in accordance with the claimed invention is the reliance on the German word “kommen” to satisfy the claim limitations requiring the claimed cooperative conversational voice user interface to receive an utterance including “one or more words that have different meanings in different contexts,” and “disambiguat[e] / [ing] an intent the user had in speaking the one or more words that have different meanings in different contexts.” *See, e.g.*, cl. 1; Petition, 14-15.

73. The Petition states that SmartKom’s “intention recognizer establishes the intended meaning of an utterance within the identified context—i.e., disambiguates an intent that the user had in speaking the one or more words that have different meanings in different contexts.” Petition, 36. However, the Petition describes the functionality of the intention recognizer using the example from SmartKom where a user request “Is there a movie with Nicole Kidman?” leads to two intention hypotheses: one where the user requests information about the current TV program, and one where the user requests information about the current cinema program. Petition, 28-32 (citing Ex. 1005, 263-265). The Petition does not assert that “Is there a movie with Nicole Kidman?” is an utterance that includes “one or more words that have the different meanings in the different contexts.”

74. Rather, the Petition cites to pages 280 and 607 of SmartKom to support its assertion that SmartKom discloses “disambiguat[ing] an intent the user had in

speaking the one or more words that have different meanings in the different contexts” based on the disclosure “that the verb “kommen” can have different meanings depending on the context. Petition, 37. However, page 607 of SmartKom merely discloses that, “[d]ue to lexical ambiguity of individual words, each SRH can be mapped to a set of possible interpretations called *conceptual representations* (CR). E.g., the German word *kommen* can be mapped to either the ontological concept MotionDirectedProcess or to the concept WatchPerceptualProcess.” Ex. 1005, 607 (italics in original). Likewise, page 280 states that “[t]he evaluation of the contextual coherence scoring in terms of its disambiguation performance meant calculating how often contextual coherence picks the appropriate reading, given an ambiguous lexicon entry such as *kommen* associated in the lexicon with both MotionDirectedProcess and WatchPerceptualProcess.” Ex. 1005, 280; Petition, 37-38. Due to the lack of explicit disclosure of the different meanings in different contexts in Petitioners’ citation to SmartKom, the Petition relies on another article authored by Porzel and cited in SmartKom (Ex. 1014) that is not part of the asserted ground to explain two example input utterances using “kommen.” Petition, 15. The Petition does not point to any other disclosure in the asserted references that discloses a word having “different meanings in different contexts” as required by the claims. For the foregoing reasons, it is my opinion that the Petition fails to demonstrate that the SmartKom + Kobsa combination renders obvious the

challenged claims because the asserted combination fails to disclose or teach a coherent system in accordance with the claims.

C. A POSITA Would Not be Motivated to Combine SmartKom and Kobsa to Reach the Claimed Invention

75. In my opinion, the SmartKom and Kobsa combination fails to render obvious the challenged claims because a POSITA would not have been motivated to combine SmartKom and Kobsa to reach the claimed invention.

76. I understand the Petition asserts that, among other reasons, a POSITA would have been motivated to combine Kobsa's teachings regarding user models and user modeling with SmartKom's dialogue system. Petition, 8. Petitioners' alleged motivation to combine SmartKom and Kobsa relies on SmartKom's mention of the "use of a user model and/or profiles/preferences in its context modeling and presentation planning," and mention of "storing preferences stated by a user during a conversation," but its failure to disclose further details about these user models. Petition, 8. I disagree that a POSITA would have been motivated to combine Kobsa with SmartKom to reach the claimed invention.

77. Specifically, as I explained above, Petitioners assert that Kobsa's disclosure of General User Modeling System which discloses a "persistent" user model and SmartKom's disclosure of a domain model, user model, or interaction model satisfies the "long-term shared knowledge functionality in the claims. As set forth above, the Petition offers no explanation regarding how a POSITA would

combine these disclosures in accordance with the claims. For example, there is no explanation as to how Kobsa's General User Modeling System would be integrated into SmartKom's dialogue system in either (or all of) the domain model or the user model, which Petitioners allege to be used in context modeling module, annotated in yellow, and/or the interaction modeling module annotated in green. Notably, the processing in these modules respectively occur before and after the intention is recognized.

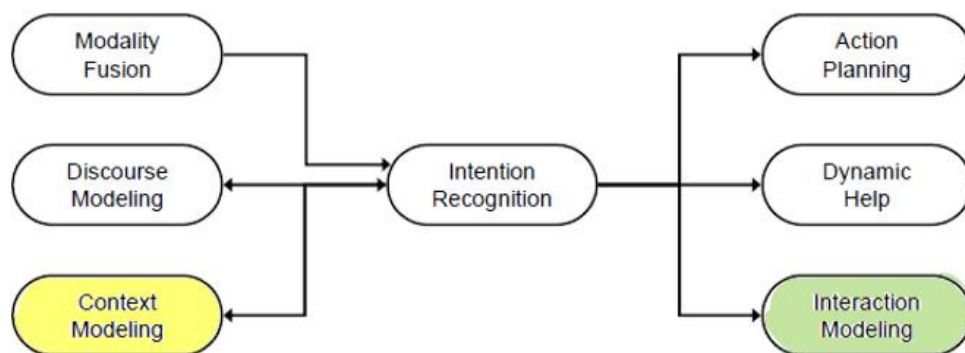


Fig. 1. Section of the multimodal interaction from the intention recognition point of view

78. It is my opinion that, even if the Petition provided an explanation, due to the differences between Kobsa and SmartKom's disclosures, a POSITA would not be motivated to combine Kobsa's General User Modeling System with SmartKom's domain model, user model, or interaction model to meet the "long-term shared knowledge" limitations. With respect to the implementation of a user model in SmartKom, SmartKom discloses that "general user model information is supplied

via external sources, e.g., via a user's *SmartCard*," while the SmartKom system only "actively monitors" "interaction preferences of the users." Ex. 1005 at 276 (emphasis in original). Accordingly, apart from interaction preferences, SmartKom discloses that there is no active monitoring of user information. An external plug-in user model with no active monitoring of user information other than interaction preferences in SmartKom's system is completely inconsistent with the "persistent" user model in Kobsa, which "records *all information* about the user inferred from his/her dialog behavior" "at the end of a dialog session." Petition, 23; Ex. 1006, 10-11.

79. Petitioners have not pointed to any disclosure in SmartKom, and I have not seen any disclosure, explaining that an external plug-in user model would or could be "persistent" or continually updated to meet the "accumulating long-term knowledge" limitations of the '681 Patent claims. Petitioners have not explained how the persistent user model of Kobsa could be incorporated into a plug-in SmartCard, for example, to be used in SmartKom's system such that the combined system would work in accordance with the claims.

80. Furthermore, as I explained above, to the extent Petitioners allege that actively monitored interaction preferences would meet these limitations, the Petition does not explain how Kobsa's user model would be incorporated with the actively monitored interaction preferences (e.g., though interaction modeling) of SmartKom

in the same combined system. For the foregoing reasons, it is my opinion that the Petition fails to demonstrate that a POSITA would be motivated to combine SmartKom and Kobsa in accordance with the claimed invention.

D. The Petition Fails to Show that the SmartKom + Kobsa Combination Discloses or Teaches “short-term shared knowledge”

81. The Challenged Claims require “accumulat[e]/[ing] short-term shared knowledge about the current conversation, wherein the short-term shared knowledge includes knowledge about the utterance received [at the voice] during the current conversation.” Ex. 1001, cl. 1, 13, 25. In my opinion, the Petition fails to show that SmartKom in combination with Kobsa discloses or renders obvious this limitation.

82. The Petition relies only on SmartKom’s disclosure for these limitations. Specifically, with respect to the portion of these limitations directed to “short-term shared knowledge,” the Petition states that the dialogue model, the situational model, and the storage of user state are types of “short-term shared knowledge” that “includes knowledge about the utterance received. To explain further, the Petition states that the dialogue model that “contains ‘what has been said by who’ and is associated with dialogical (discussion/conversation) context” includes the dialogue history, or the discourse state, which is a type of “short-term shared knowledge”. Pet. at 18-19. The Petition further identifies the situational model, which records “time, place” of an utterance as “short-term shared knowledge.” Pet. at 19. The

Petition further identifies SmartKom’s disclosure of storage of the user state, such as “hesitant,” is a type of short-term shared knowledge. Pet. at 19.

83. For the limitations requiring “accumulate[ing]/[e] short-term shared knowledge about the current conversation” and that the “short-term shared knowledge includes knowledge about the utterance received [at the voice] during the current conversation,” the Petition identifies two alternative ways that SmartKom allegedly discloses these limitations.

84. First, the Petition identifies the discourse model of SmartKom, and SmartKom’s disclosure that the discourse model “is dynamically updated as system output progresses.” Petition, 21 (quoting Ex. 1005, 61). However, I have not seen any disclosure in SmartKom, and Petitioners do not identify any disclosure, indicating that the dialog history in the discourse model is about the current conversation with the user. Furthermore, dynamically updating the discourse model as system output progresses fails to teach or suggest any distinction between utterances received during the current conversation with the user, as required by the claims, and any prior conversation, for example. That the dialog history updated in a manner characterized by constant change as system output progress does not teach a temporal aspect that the knowledge about what has been said is related to the current conversation. Mr. Lipoff’s testimony regarding this alternative does not change my opinion. The only additional support that Mr. Lipoff offers is an example

of the discourse state being “dynamically updated as system output progresses,” wherein “the intention recognition component selects an intention hypothesis for the current utterance . . . is incorporated into the discourse history representing a user turn.” Ex. 1003, ¶ 103. I disagree that this example demonstrates the claim limitation because Mr. Lipoff and Petitioners have not identified any disclosure from SmartKom that demonstrates the user and system turns identified by Mr. Lipoff are related to “the current conversation” as required by the claims. Petitioners have not identified any distinction between the discourse state as it relates to the current conversation as opposed to prior conversations, for example, in SmartKom.

85. Petitioners cite to a portion of SmartKom that describes the discourse state as being “dynamically update as system output progresses,” but ignore that SmartKom also teaches a process of overlay for managing the discourse state, which is an “operation capable of overwriting parts of the old structure with the new information and at the same time keeping the information still consistent with the new information.” Ex. 1005, 246; Petition, 21. This process is based on default unification for “overwriting old, contextual information—*background*—with new information—*covering*—thereby keeping as much consistent information as possible.” Ex. 1005, 246. Petitioners do not address the overlay process used in the implementation of discourse processing described in SmartKom, which underscores Petitioners’ failure to identify “accumulate[ing]/[e] short-term shared knowledge

about the current conversation” and that the “short-term shared knowledge includes knowledge about the utterance received [at the voice] during the current conversation,” by their identification of a dynamically updated discourse state.

86. The second alternative identified by the Petition as the disclosure of accumulate[ing]/[e] short-term shared knowledge about the current conversation” and that the “short-term shared knowledge includes knowledge about the utterance received [at the voice] during the current conversation,” as required by claims is SmartKom’s disclosure of a user state lattice wherein the prosody module uses speech signals and word lattices from the speech recognizer and generates a user state lattice. Pet., 22. Again, however, I disagree that this disclosure includes any limitation that the user state lattice is about the current conversation for similar reasons I discussed for the first alternative. Namely, neither Petitioners nor Mr. Lipoff point to any disclosure in SmartKom that user state information is received about the current conversation during the current conversation.

VIII. GROUNDS 2-4

87. The remaining Ground 2-4 of the Petition rely on combinations of Barbara (Ex. 1007) and Ross (Ex. 1008). In my opinion, the Petition fails to show that the asserted prior art in Grounds 2-4 render obvious the challenged claims.

A. Overview of the Asserted References

1. *Barbara* (Ex. 1007)

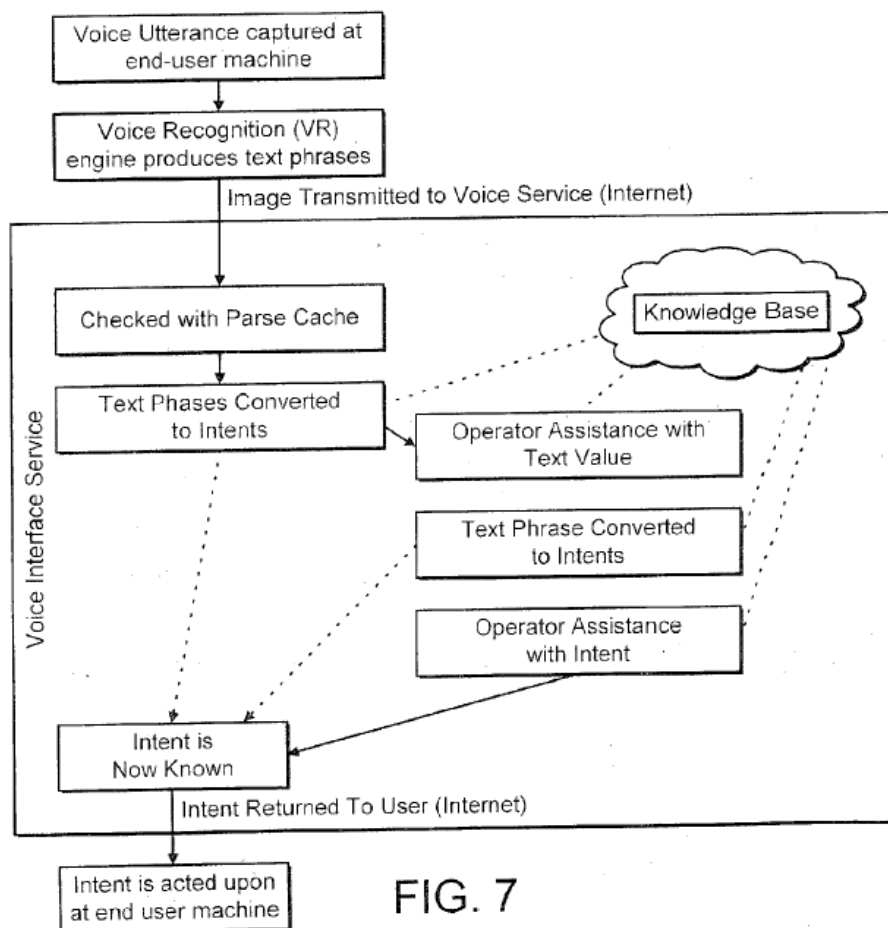
88. The portions of *Barbara* relied upon by Petitioners are directed to systems and methods for creating a voice interface where the end system responds to spoken natural language commands. Ex. 1007, ¶[0082]. *Barbara* discloses that the “electronic version of the spoken word is checked for accuracy, corrected and interpreted to determine what the user wanted.” Ex. 1007, ¶[0082]. In *Barbara*, when an utterance has been detected, “a request packet is then generated, the request comprising the best guesses of the voice recognition engine as to what was said, with some indications of their likelihood and also an audio file of the utterance.” Ex. 1007, ¶[0086]. “The best guesses of the voice recognition engine as to what was said are evaluated against an information database(s),” which is “built up over time, by recording personal details or preferences from what has been said, or from information directly entered into the system by the user.” (Ex. 1007, ¶[0088]). The best guesses for the possible requests are also evaluated against a knowledge base, where “the correct utterance text and the correct intention are stored.” Ex. 1007, ¶[0106].

89. *Barbara* discloses that “[w]hen each utterance is received at the server, it is evaluated into a set of possible requests that the user may have actually meant,” and a “combined likelihood for each possible request is heuristically calculated.” Ex.

1007, ¶[0090]. Part of this process includes the information database being “searched for words that are similar to or the same as those interpreted,” and “if anything is identified, then a message is sent . . . to the server identifying the word from the information database and providing an indication as to its meaning and the confidence of that interpretation.” Ex. 1007, ¶[0089].

90. Barbara’s “automatic recognition process is to consider every sub-phrase from the sentence in order of increasing length,” which begins by “mapping to intentions all sub-phrases,” in increasing length which is “continued . . . until a set of possible intentions is determined for the whole phrase.” Ex. 1007, ¶[0118]. Then, “[c]onfidences are heuristically calculated for each possible intention, by considering the mapping rules used and their ‘success’ in the history” of the system. Ex. 1007, ¶[0118]. Importantly, the knowledge based is continually used to calculate rule confidences because the rule confidences are periodically calculated “based on analysis of the knowledge base.” Ex. 1007, ¶[0118].

91. Importantly, Barbara’s recognition process is fundamentally based on human operators. Barbara’s reliance on human operators exists in two tiers: (i) there is a backstop for speech recognition where the human operator chooses the correct text of the spoken word, and (ii) there is backstop for determining the intent of the user’s utterance when the system cannot deduce it, even though the system knows was spoken. Barbara’s reliance on human operators is shown below in Figure 7.



Ex. 1007, Fig. 7

92. As shown in Figure 7, Barbara's system requires operator assistance with determining both the correct text value of the utterance, and the intent.. However, where utterances have failed to meet a criteria for automatic acceptance, they are replayed for a human operator who chooses the correct textual interpretation. Once the textual interpretation has been deduced with the assistance of the human operator, the interpretation process begins. Barbara discloses that, in

cases where it is not possible to automatically deduce what the user wanted to do despite knowing what they said, the message is sent to a second human operator who hears what was said and determines what the utterance means. Ex. 1007, ¶ [0104]-[0105]. Thus, Barbara's process is fundamentally reliant on human assistance.

2. Ross (Ex. 1008)

93. Ross (Ex. 1008) is directed to a system that uses a specific domain model for a particular application to associate a particular syntax such that, when the system interacts with a user, there is a more natural interaction and better consistency between the input and output. Ex. 1008, ¶[1005].

94. The disclosure especially relevant to my opinions below relied on by Petitioners is Ross's disclosure of a conversational record. Ex. 1008, ¶[0058]. Ross's conversation manager in Figure 3 below stores conversational record 60. Ex. 1008, ¶[0033].

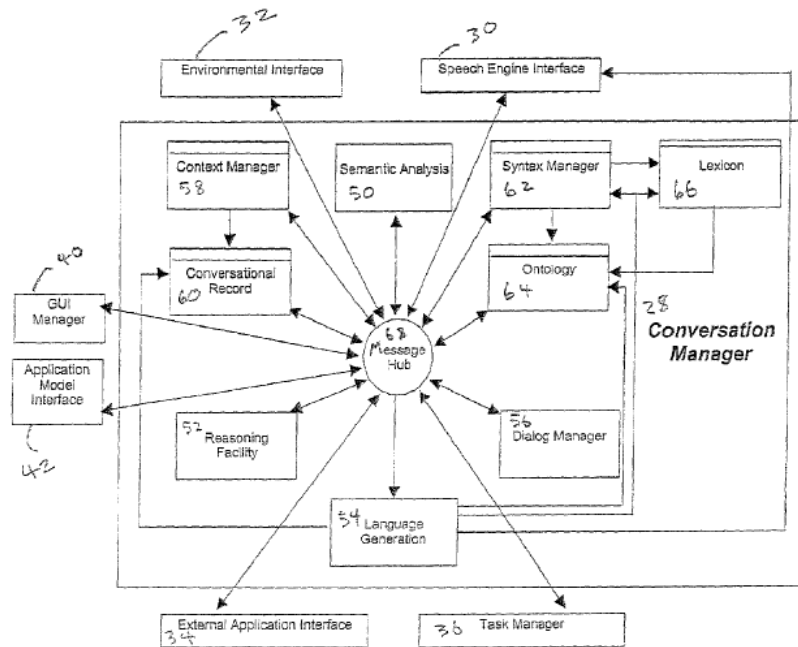


Fig. 3

Ex. 1008, Fig. 3

95. For Grounds 2-4, I understand that Petitioners rely on Ross’s architecture of the conversation manager and “data organization” including the conversational record, because of Ross’s disclosure of purging information from the “conversational record when it is no longer relevant to active goals and after some predetermined time has elapsed.” Ex. 1003, ¶[0057].

B. A POSITA Would Not Be Motivated to Combine Barbara and Ross

96. I understand that the Petition alleges that a POSITA would have been motivated to combine Barbara and Ross. Specifically, the Petition alleges that “a

POSITA would have been motivated to combine Ross's teachings regarding the architecture of a conversation manager with Barbara's voice interface system. Petition, 65. The Petition further asserts that, while Barbara discloses use of an "information database" and "knowledge base" to interpret utterances, a POSITA would have been motivated to combine Barbara's databases with Ross's conversations records for storing a dialog history because Barbara "provides limited detail regarding managing the information" in its data stores. Petition, 65-66. Thus, Petitioners assert that "[i]mplementing Ross's teachings regarding purging information from the 'conversational record' allows Barbara's dialog history to be expired and long-term data to be retained." Petition, 66. I disagree with this characterization of the references and this motivation because Ross only discloses purging the conversational record, and does not teach or suggest retaining long-term data.

97. Furthermore, Petitioners allege that a POSITA would have been motivated to combine Barbara and Ross because, while Barbara discloses use of an information database and knowledge base to interpret utterances, it does not provide extensive details about managing the information in those data stores, whereas Ross teaches details regarding storage of dialog history because of its disclosure of purging the conversational record. Petition, 65-66. I disagree with the reasons provided by Petitioners that POSITA would have been motivated to combine

Barbara and Ross. In my opinion, a POSITA would not have been motivated to combine the disclosures of Barbara and Ross to achieve the claimed invention.

98. I disagree with Petitioners assertion that a POSITA would modify Barbara's knowledge and information databases using Ross's organization and architecture that includes purging information from the conversational record when it is no longer relevant to active goals and after some predetermined time. This discarding of information is entirely inconsistent with the automatic recognition process in Barbara, the successful function of which is reliant on the information database and knowledge base being built up overtime and continually maintained.

99. Specifically, as I explained above, the knowledge base stores "the correct utterance text and the correct intention." Ex. 1007, ¶[0106]. Petitioners rely on Barbara's recognition process for the majority of the claim limitations, and that process works by mapping the intentions of each individual sub-phrase and word of an utterance, and calculating the corresponding confidences for accurate recognition. These confidences in Barbara are calculated based on an analysis of the knowledge base. Ex. 1008, ¶[0118]. Petitioners specifically rely on Barbara's knowledge base and information database as allegedly disclosing or teaching the limitations of the Challenged Claims directed to short-term shared knowledge. Petition at 72.

100. In Barbara, the history of correct utterance text and correct intentions are continually used to calculate the confidences. Thus, the continual maintenance

of the history of correct utterance text and correct intentions in the knowledge base is vital to the operation of Barbara's recognition process. That is, in order for Barbara's system to work, the knowledge base must be continually maintained and be as "broadly applicable as possible." Ex. 1007, ¶¶0110].

101. Likewise, the information database in Barbara "is built up over time, by recording personal details from what has been said[.]" Ex. 1007, ¶¶0088]. During Barbara's recognition process, the information database is searched for words that are similar to or the same. Ex. 1007, ¶¶0089]. Accordingly, the functionality of Barbara's system that Petitioners rely on is dependent on the information database maintaining the history of what has been said so that future interpretations are successful.

102. In contrast, Ross's conversational record indexes each utterance with the results of its semantic analysis and eventually purges it when it is no longer relevant and after a predetermined period of elapsed time. Ex. 1008, ¶¶0057].

103. I note that Petitioners appear to identify both Barbara's knowledge base and information database as satisfying the "short-term shared knowledge" limitations of the claims, Petition at 72. However, Petitioners only allege that Barbara's knowledge based would be modified by Ross's data organization including the conversational record, and fails to make the same assertion for the information database. Petition at 72-73. In any case, I disagree that a POSITA would

be motivated to modify Barbara's information database or knowledge base using Ross's conversational record.

104. Petitioners allege that a POSITA would have modified Barbara's knowledge base using Ross's data organization to make the knowledge base short-term storage that "is eventually purged from the conversational record when it is no longer relevant. Petition, 72-73. I disagree, because periodically purging the information stored in the knowledge base is entirely incompatible with the function and purpose of the knowledge base in Barbara's recognition system.

105. If Barbara's knowledge base was periodically purged, as Petitioners assert with this combination, Barbara's recognition process would not be able to successfully calculate the confidences of each interpreted word because there would be no history in the knowledge base to calculate these confidences. Barbara's knowledge base has to continually maintain all correct interpretations in order to continually be used to calculate the confidence scores for each word. Even though I understand Petitioners have not made the specific assertion to modify the information database with Ross's conversational record, that modification would not work either because the history, "built up over time" based on what has been said is also analyzed in order to interpret utterances received by Barbara's system. Thus, a POSITA would not be motivated to modify either the information database or the knowledge base using Ross's period purge of information because Barbara's system

simply would not work if modified in the way that Petitioners suggest. Thus, in my opinion, a POSITA would not be motivated to combine these references as Petitioners suggest to achieve the claimed invention.

106. Moreover, the Petition asserts that Barabra and Ross are analogous art to the '681 Patent pertaining to the same endeavor. I disagree with the Petition's assertion that a POSITA would have been motivated to combine Barbara and Ross as analogous references. Barbara, on the one hand, is directed to a communication system where the voice input from a user terminal is routed to a plurality of operator terminals manned by human operators that perform manual interpretation of the incoming information from the user terminals. Ex. 1007, ¶¶[0083]-[0085]. Ross, on the other hand, is directed to language generation method that performs its work in the context of a domain model for a particular application so that dialog between the user and the system is more natural. Ex. 1008, ¶¶[0004]-[0005]. Not only are the references directed to different domains, and solutions to different problems, but unlike Barbara, Ross's system is not reliant on manual interpretation by human operators. In my opinion, these differences would deter a POSITA from combining these references in the manner asserted by the Petition.

**C. The Barbara + Ross Combinations Fails to Disclose or Teach
“determining an intended meaning” because Barbara is based on
human assistance**

107. The Petition asserts that the Barbara and Ross combination in Grounds 2-4 render obvious the claim limitations requiring “determining an intended meaning.” I disagree.

108. The design of Barbara’s system that performs intention recognition is fundamentally based on human assistance, which is inconsistent with the claims. Ex. 1007, Fig. 7, ¶¶ [0100], [0104]. Petitioners rely on Barbara’s voice interface system for receiving, interpreting, and correcting incoming information in combination with Ross and additional secondary references as allegedly rendering obvious the asserted claims. Petition at 67. However, as I explained above in Section VII.A above, Barbara’s voice interface system is designed such that human operators, not the system itself, is ultimately used for the utterance text recognition and the intention recognition.

109. I have reviewed the intrinsic record of the ’681 Patent, including its specification and file history. Based on my analysis, it is my opinion that a system that uses human assistance “determine an intended meaning” of an utterance does not meet the challenged claims.

110. The intrinsic record of the ’681 Patent demonstrates that the claimed cooperative conversational voice interface system does not use human operators for

assistance in determining the intended meaning. Instead, the claims are directed to creating a Human-to-Machine interface. Ex. 1001, 13:10-27.

111. For example, certain embodiments of the '681 Patent invention explain that context domain agents and an intelligent hypothesis builder *of the system* are employed for context determination and determination of intended meaning. *See e.g., id.*, 12:19-56, 15:35-16:23.

112. The prosecution history confirms my understanding. During prosecution, references were cited against the '681 Patent application that included human assistance for determining the intended meaning of an utterance. Ex. 1002, 235. In response to the rejection, the applicant argued that the cited reference failed to “disclose, teach, or suggest that the system determines an ‘intended meaning’ for an utterance” because the reference described a system that asked a caller about whether a call entity recognized in an utterance is incorrect instead of attempting to use trained speech recognition algorithms to *independently disambiguate* the intended meaning for the utterance. Ex. 1002, 235. Based on these statements, it is my opinion that human operator assistance in determining the intended meaning of the utterance is inconsistent with the claim limitations.

113. Accordingly, in my opinion, Grounds 2-4 based on the combination of Barbara and Ross fail to render obvious the challenged claims.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on the information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on August 20, 2025 in San Jose, CA.

A handwritten signature in black ink, appearing to be 'H. Homayoun', written over a horizontal line.

Houman Homayoun, Ph.D.