

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Comcast Cable Communications, LLC
Petitioner

v.

Promptu Systems Corporation
Patent Owner

Patent No. 7,047,196

DECLARATION OF CHRISTOPHER SCHMANDT
REGARDING U.S. PATENT NO. 7,047,196

(Submitted in Support of Petition 1 of 2)

TABLE OF CONTENTS

I.	Background and Qualifications.....	1
II.	Materials Considered	4
III.	Legal Standards for Patentability	6
IV.	Background on Speech Recognition in Cable Television Systems	12
	A. Server-Based Speech Recognition	12
	B. Server-Based Speech Recognition in Cable Television Systems	17
V.	The '196 Patent	22
	A. Overview of the '196 Patent	22
	B. Prosecution History of the '196 Patent.....	25
	C. Claims of the '196 Patent.....	27
	D. Effective Filing Date of the '196 Patent.....	41
	E. Level of Ordinary Skill in the Art.....	42
VI.	Construction of Terms Used in the '196 Patent.....	45
	A. “wireline node” (claims 1, 14, 27, 53)	45
	B. “back channel” (claims 1, 2, 14, 27, 28, 53).....	46
	C. “partitioning said received back channel into a multiplicity of [said] received identified speech channels” (claims 1, 2, 27, 28)	47
VII.	Overview of Prior Art References.....	51
	A. Overview of Murdock	51
	B. Overview of Julia	64
	C. Overview of Nazarathy.....	67
	D. Overview of Quigley	69
	E. Overview of Banker.....	72
	F. Overview of Gordon.....	73
VIII.	The Challenged Claims Are Rendered Obvious by Murdock Alone Or In Combination with Nazarathy, Quigley, Banker, Or Gordon.....	74
	A. Every Element of the Challenged Claims Is Disclosed by Murdock Alone or Combined with Nazarathy, Quigley, Banker, or Gordon	75
	1. Claim 1	75

TABLE OF CONTENTS
(continued)

2.	Claim 2	89
3.	Claim 4	96
4.	Claim 5	97
5.	Claim 6	103
6.	Claim 12	110
7.	Claim 13	112
8.	Claim 27	113
9.	Claim 28	127
10.	Claim 30	131
11.	Claim 31	132
12.	Claim 32	134
13.	Claim 38	139
14.	Claim 39	140
15.	Claim 40	141
16.	Claim 41	144
17.	Claim 42	147
B.	A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Murdock with the Teachings of Nazarathy and Quigley	149
C.	A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Murdock with the Teachings of Either Banker or Gordon	154
IX.	The Challenged Claims Are Rendered Obvious by Julia In Combination with Nazarathy, Quigley, Banker, or Gordon.....	157
A.	Every Element of the Challenged Claims Is Disclosed by Julia Combined with Nazarathy, Quigley, Banker, or Gordon.....	159
1.	Claim 1	159
2.	Claim 2	172
3.	Claim 4	177

TABLE OF CONTENTS
(continued)

4.	Claim 5	177
5.	Claim 6	182
6.	Claim 12	188
7.	Claim 13	190
8.	Claim 27	192
9.	Claim 28	203
10.	Claim 30	207
11.	Claim 31	208
12.	Claim 32	210
13.	Claim 38	214
14.	Claim 39	215
15.	Claim 40	216
16.	Claim 41	218
17.	Claim 42	220
B.	A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Julia with the Teaching of Nazarathy and Quigley	222
C.	A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Julia with the Teachings of Either Banker or Gordon	227
X.	Conclusion.....	230

I, Christopher Schmandt, declare as follows:

1. I make this declaration based upon my own personal knowledge and, if called upon to testify, would testify competently to the matters stated herein.

2. I have been asked by Comcast Cable Communications, LLP to provide technical assistance in connection with the *inter partes* review of U.S. Patent No. 7,047,196 (which I will refer to as the “’196 Patent”). This declaration is a statement of my opinions on issues related to the patentability of claims 1-2, 4-6, 12-13, 27-28, 30-32, and 38-42 (which I will refer to as the “challenged claims”) of the ’196 Patent.

3. My compensation is not based on the content of my opinions or the resolution of this matter.

I. BACKGROUND AND QUALIFICATIONS

4. In formulating my opinions, I have relied on my knowledge, training, and experience in the relevant field, which I will summarize briefly.

5. I am currently employed as a Principal Research Scientist at the Media Laboratory at the Massachusetts Institute of Technology (MIT). In that role I also serve as faculty for the MIT Media Arts and Sciences academic program. I have more than 35 years of experience in the field of Media Technology, and I was a founder of the MIT Media Laboratory.

6. I received my B.S. degree in Electrical Engineering and Computer

Science from MIT in 1978, and my M.S. in Visual Studies (Computer Graphics) also from MIT. I have been employed at MIT since 1980, initially at the Architecture Machine Group, which was an early computer graphics research lab. In 1985, I helped found the Media Laboratory and continue to work there now. I currently run a research group titled “Living Mobile.” My research spans distributed communication and collaborative systems, with an emphasis on multimedia and user interfaces. I have over 70 published conference and journal papers and one book in these fields.

7. For the first fifteen years of my career, my research emphasized speech recognition and speech user interfaces. I built the first conversational computer system utilizing speech recognition and synthesis (“Put That There”) starting in 1980. I continued to innovate speech user interfaces using recognition, text-to-speech synthesis, and recorded audio in a wide variety of projects. I was later awarded membership to the Association for Computing Machinery (ACM) Computer Human Interface (CHI) Academy specifically for those years of work pioneering speech user interfaces.

8. In the 1980s and 1990s a major research theme of the Media Laboratory was the convergence of networking and digital media, including the delivery of interactive information and entertainment systems to the home, by both cable and wireless. Distributed multimedia systems required both networking

means to deliver content and software and user interfaces to afford interactivity. In addition to conventional computer networking, we also had in house a broadband cable television system capable of running hybrid video and data channels. We developed software to implement media servers that could act as components in a cable system head end, and interactive client-side software and user interfaces to manage the content for consumption by the consumer.

9. In the course of my research I built a number of speech recognition client/server distributed systems, with the first being in 1985. Much of the initial motivation for a server architecture was that speech recognition required expensive digital signal processing hardware that we could not afford to put on each computer, so a central server with the required hardware was used. Later versions of the speech recognition server architecture allowed certain computers to perform specialized tasks serving a number of client computers providing voice user interfaces, either on screens or over telephone connections.

10. Because of my early work with distributed speech systems, I served for several years in the mid-1990s on a working group on the impact of multimedia systems on the Internet reporting to the Internet Engineering Task Force (IETF) and later the Internet Activities Board (IAB). This work impacted emerging standards such as Session Initiation Protocol (SIP).

11. In my faculty position I teach courses and directly supervise student

research and theses at the Bachelors, Masters, and PhD levels. I oversee the Masters and PhD thesis programs for the entire Media Arts and Sciences academic program. I also have served on the Media Laboratory intellectual property committee for many years.

12. Based on the above experience and qualifications, I have a solid understanding of the knowledge and perspective of a person of ordinary skill in this technical field since at least 1990.

II. MATERIALS CONSIDERED

13. I have reviewed and considered the following documents, among others, in connection with my analysis of the '196 Patent:

- the '196 Patent (Ex. 1001);
- the prosecution history of U.S. Patent Application No. 09/785,375, which led to the issuance of the '196 Patent (“File History”) (Ex. 1002);
- Provisional Application No. 60/210,440 (“Provisional Appl.”) (Ex. 1003);
- “Speech Recognition by Machine: A Review” by D. Raj Reddy, Proceedings of the IEEE, Vol. 64, No. 4, April 1976 (“Reddy”) (Ex. 1004);
- “Multimodal User Interfaces in the Open Agent Architecture” by D.

Moran, et al., Proceedings of the 2nd International Conference on Intelligent User Interfaces (IUI '97), pp. 61-68, 1997 (“Moran”) (Ex. 1005);

- “Quantization of Cepstral Parameters for Speech Recognition over the World Wide Web” by V. Digalakis, et al., IEEE Journal on Selected Areas in Communications, Vol. 17, No. 1, Jan. 1999 (“Digalakis”) (Ex. 1006);
- “Enabling New Speech Driven Services for Mobile Devices: An Overview of the ETSI Standards Activities for Distributed Speech Recognition Front-Ends” by D. Pearce, AVIOS 2000: The Speech Applications Conference, 2000 (“Pearce”) (Ex. 1007);
- “IBM Technical Disclosure Bulletin: Speech Recognition Methods for Controlling Cable Television,” Vol. 38, No. 08, August 1995 (“IBM Bulletin”) (Ex. 1008);
- U.S. Patent No. 6,345,389 (“Dureau”) (Ex. 1009);
- U.S. Patent No. 7,013,283 (“Murdock”) (Ex. 1010);
- Provisional Application No. 60/166,010 (“Murdock Provisional”) (Ex. 1011);
- U.S. Patent No. 6,513,063 (“Julia”) (Ex. 1012);
- U.S. Patent No. 6,490,727 (“Nazarathy”) (Ex. 1013);

- U.S. Patent No. 6,650,624 (“Quigley”) (Ex. 1014).
- U.S. Patent No. 5,477,262 (“Banker”) (Ex. 1015);
- U.S. Patent No. 6,314,573 (“Gordon”) (Ex. 1016); and
- Plaintiff Promptu Systems Corporation’s Disclosure of Asserted Claims and Preliminary Infringement Contentions, dated December 11, 2017 (“Infringement Contentions”) (Ex. 1017).

14. In addition, my opinions are based on my years of experience in the field of interactive television and other multimedia systems.

III. LEGAL STANDARDS FOR PATENTABILITY

15. My opinions are also formed by my understanding of the relevant law. I am not an attorney. For purposes of this declaration, I have been informed about certain aspects of the law as it relates to my opinions.

16. I understand that for an invention claimed in a patent to be found patentable, it must be (among other things) new and not obvious based on what was known before the invention was made.

17. I understand that the information that is used to evaluate whether an invention was new and not obvious when made is generally referred to as “prior art.” I understand that the prior art includes all patents and printed publications that existed before the earliest filing date of the patent (i.e., the “effective filing date”). This includes foreign language material. I also understand that a patent is

prior art if it was filed before the effective filing date of the claimed invention and that a printed publication is prior art if it was publicly available before the effective filing date.

18. I understand that in this *inter partes* review proceeding, Comcast has the burden of proving that the challenged claims of the '196 Patent are unpatentable in light of the prior art by a preponderance of the evidence. I understand that a preponderance of the evidence is evidence sufficient to show that a fact is more likely true than not true.

19. I understand that unexpired patent claims in an *inter partes* review are given the broadest reasonable interpretation in light of the patent specification. After the claims are construed, they are then compared to the prior art.

20. I understand that in this *inter partes* review proceeding, the information that may be evaluated is limited to patents and printed publications. My analysis, which is set out in detail below, compares the challenged claims of the '196 Patent to patents and printed publications that are prior art to the claims.

21. I understand that prior art can render the claim “obvious” to a person of ordinary skill in the art. My understanding of this legal standard is set out below.

22. I understand that prior art can render a patent claim unpatentable where subject matter that falls within the scope of the claim would have been

obvious to a person of ordinary skill in the art. I understand that the following standards govern the determination of whether a patent claim is rendered “obvious” in light of the prior art. I have applied these standards in my evaluation of whether the challenged claims of the ’196 Patent are obvious in light of the prior art.

23. I understand that a claimed invention is not patentable if it would have been obvious to a person of ordinary skill in the field of the invention at the time the invention was made. The person of ordinary skill in the art is a hypothetical person who is presumed to have known the relevant art at the time of the invention. Even if all the requirements of a claim were not found in a single prior art reference, the claim is not patentable if the differences between the subject matter in the prior art and the subject matter in the claim would have been obvious to a person of ordinary skill in the art at the time the application was filed. Prior art disclosing a method or device that falls within the scope of a claim can render that claim obvious even if other, different methods or devices might also fall within the scope of the claim.

24. I understand that a determination as to whether a claim would have been obvious should be based on four factors (though not necessarily in the following order): (i) the level of ordinary skill in the art at the time the application was filed; (ii) the scope and content of the prior art; (iii) the differences between

the claimed invention and the prior art; and (iv) any “objective factors” indicating obviousness or non-obviousness that may exist in a particular case.

25. I understand that an obviousness analysis should not be based on hindsight, but must be done using the perspective of a person of ordinary skill in the relevant art as of the effective filing date of the patent claim.

26. I understand the objective factors indicating obviousness or non-obviousness may include: commercial success of products covered by the patent claims; a long-felt but unaddressed need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; expressions of surprise by experts and those skilled in the art at the making of the invention; and the patentee having proceeded contrary to the accepted wisdom of the prior art. I also understand that any of this evidence must be specifically connected to the claimed invention rather than associated with the prior art or with marketing or other efforts to promote an invention. I am not presently aware of any evidence of such objective factors suggesting the claims of the '196 Patent are non-obvious. Should the Patent Owner submit evidence purportedly showing such objective factors, I reserve the right to consider the evidence and respond to it.

27. I understand that the teachings of two or more prior art references may be combined in the manner disclosed in the claim if such a combination would have been obvious to one having ordinary skill in the art. In determining whether a

combination would have been obvious, the following exemplary rationales may support a conclusion of obviousness:

- combining prior art elements according to known methods to yield predictable results;
- simple substitution of one known element for another to obtain predictable results;
- use of a known technique to improve similar devices (methods, or products) in the same way;
- applying a known technique to a known device (method, or product) ready for improvement to yield predictable results;
- “obvious to try” – choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success;
- known work in one field of endeavor may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations are predictable to one of ordinary skill in the art;
- some teaching, suggestion, or motivation in the prior art that would have led one of ordinary skill to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention; and

- common sense.

28. I understand that the obviousness analysis need not seek out precise teachings directed to the specific subject matter of the challenged claim, but instead can take account of the ordinary innovation and experimentation in the relevant field that does no more than yield predictable results.

29. I understand that, in assessing whether there was an apparent reason to modify or combine known elements as claimed, it may be necessary to look to interrelated teachings of multiple patents and printed publications, the effects of commercial demands, and the background knowledge of a person of ordinary skill in the art. I further understand that any motivation that would have applied to a person of ordinary skill in the art, including motivation from common sense or derived from the problem to be solved, is sufficient to explain why references would have been combined.

30. I understand that modifications and combinations suggested by common sense are important and should be considered. Common sense suggests that familiar items can have obvious uses beyond the particular application being described in a prior art reference, that if something can be done once it would be obvious to do it multiple times, and that in many cases a person of ordinary skill in the art can fit the teachings of multiple patents together in an obvious manner to address a particular problem. Further, the prior art does not need to be directed to

solving the same problem that is addressed in the patent.

31. I understand that a person of ordinary skill in the art is also a person of ordinary creativity. In many fields, it may be that there is little discussion of obvious techniques, modifications, and combinations, and it may be the case that market demand, rather than scientific research or literature, will drive a new design. When there is market pressure or design need to solve a particular problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has a good reason to employ the known options. If this leads to the expected success, then it is likely the product of ordinary skill and common sense as opposed to patentable innovation. I understand that if a combination was obvious to try, that may show that it was obvious and therefore unpatentable. That a particular combination of prior art elements was obvious to try suggests that the combination was obvious even if no one made the combination.

IV. BACKGROUND ON SPEECH RECOGNITION IN CABLE TELEVISION SYSTEMS

32. The '196 Patent describes distributed speech recognition over a cable television system in which spoken user input is captured at home (or wherever content is being viewed) and transmitted to the cable headend for the actual speech recognition processing.

A. Server-Based Speech Recognition

33. Speech recognition became a serious focus of research during the

1970s. Some lead research groups were at MIT, Carnegie Mellon University (CMU), Bolt, Beranek and Newman (BBN), and the Stanford Research Institute (SRI), among others. Initial systems could recognize single words (isolated word recognition), but this soon progressed to connected or continuous speech recognition of whole sentences. A paper authored by D. Raj Reddy, who was at the Computer Science Department at CMU, entitled “Speech Recognition by Machine: A Review,” provided an overview of the already advanced state of computer speech recognition systems in 1976. Reddy (Ex. 1004). That paper provided a survey of recent developments in speech recognition and identified a number of working speech recognition systems. Reddy at 501, 506-508 (discussing “word recognition systems”), 508-511 (discussing “connected speech recognition” systems including several developed at CMU, MIT, and IBM), and 511-515 (discussing “speech understanding systems” then under development). As noted in the Reddy paper, “performance [of speech recognition systems] depends on the size and speed of the computer and the accuracy of the algorithm used.” Reddy at 503. Thus, generally speaking, a more powerful computer, such as a mainframe or server computer, would enable faster and more accurate speech recognition.

34. My research lab, the Architecture Machine Group at MIT, acquired one of the first commercial connected speech recognizers (i.e., a speech recognizer

that does not require each word to be spoken in isolation) from NEC Corp. in 1979, and I used it extensively in my own research. Within a few years the personal computer became a popular platform, after the introduction of the IBM PC, and somewhat less powerful speech recognizers were being sold to run on an expansion board that could be added inside a PC. By the early 1990s general purpose computer workstations were capable of running certain forms of speech recognition in real time without special purpose hardware.

35. Speech recognition systems made great progress in the 1990s, and by the mid-1990s various groups were building distributed speech recognition systems (i.e., capturing speech on one device and performing speech recognition processing on a different local or remote device such as in a client/server architecture). This was spurred in large part by DARPA (the Defense Advanced Research Projects Agency), which was the primary source of funding for speech recognition research in the United States at that time. DARPA funded a number of different research groups and was always interested in direct quantitative comparisons between the various groups it was supporting. This led to a directive to build distributed systems so components between research groups could be more easily mixed and matched. For example, the Spoken Language Systems Group at MIT modified their well-known Galaxy speech recognition system into a distributed system, where recognition and other aspects of natural language processing could take

place across one or more distributed computers. The new system, called Galaxy II, was made available to the research community around 1996. Galaxy II went on to be applied in a variety of information-retrieval contexts, such as air travel ticket purchasing and weather information systems. SRI developed a similar distributed architecture speech recognition system with DARPA funding as well. In the SRI system, called Gemini, “[o]nly the primary user interface agents need run on the local computer” and the “preferred speech recognizer” would run on a UNIX workstation. Moran at 61, 65 (Ex. 1005). By the mid-to-late-1990s, it was commonplace to implement speech recognition in a distributed (e.g., client/server) architecture.

36. With the growing popularity of mobile telephones and of the World Wide Web in the 1990s, researchers and developers in the computer and telecommunications industries realized the value in distributed speech recognition as well. Neither early mobile phones nor consumer grade PCs with web browsers were powerful enough for computationally intensive speech recognition, so a server-based architecture was an obvious approach for providing robust speech recognition user interfaces. In fact, by 1995, an industry working group called the Aurora Project, which included companies such as Alcatel, Ericsson, IBM, Matra, Nokia, Nortel, and Siemens, was developing a global standard for server-based speech recognition over wireless channels. Digalakis at 84 (Ex. 1006). A later

paper published by the Aurora Project in 2000 summarizing the ETSI (European Telecommunications Standard Institute) standard for Distributed Speech Recognition Front-Ends reflected one of the primary motivations for a distributed speech recognition model: client devices generally lacked processing power for more complex speech recognition and required that speech recognition be done by a more powerful server computer:

As has already happened in the wireline world, the trend to ever-increasing use of data communication is spreading to the mobile wireless world. As part of this, people want the ability to access information while on the move and the technologies to enable them to do this are now starting to be deployed. The small portable devices that will be used to access these data services cry out for improved user interfaces using speech input. At present, however, the complexity of medium and large vocabulary speech recognition systems are beyond the memory and computational resources of such devices.

Centralised servers can share the computational burden between users and enable the easy upgrade of technologies and services provided.

Pearce at 1 (Ex. 1007). The fact that industry standards for server-based speech recognition were being developed and implemented in the 1990s confirms that the concept was well-known by that time—industry standards by their nature trail behind research and development efforts and adopt technologies from among existing, proven options.

B. Server-Based Speech Recognition in Cable Television Systems

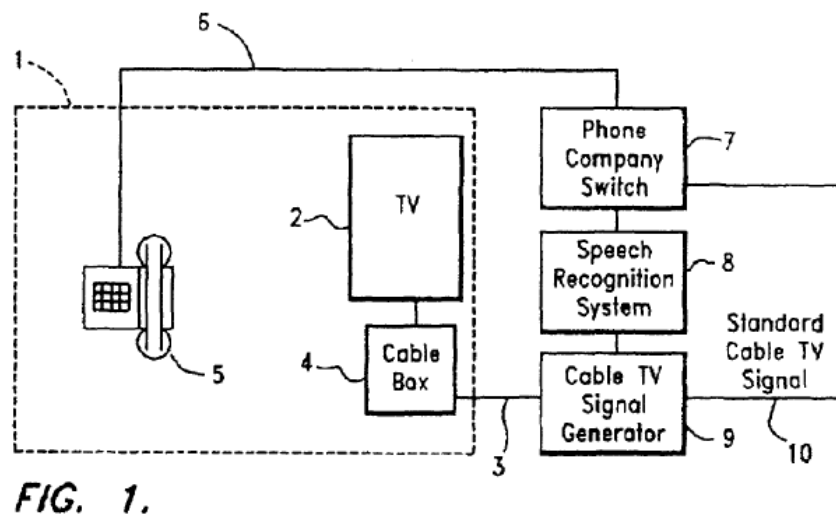
37. Speech recognition was also explored as an option for interacting with a cable television system by the mid-1990s. As cable operators acquired the technical capability to serve video content on demand, they sought complementary technologies to allow consumers to easily name a desired movie or program to watch. Speech recognition was therefore desirable as a consumer-friendly user interface for the home market to access emerging cable offerings such as pay-per-view (PPV), video-on-demand (VOD), and other new information access markets.

38. A cable television system is so named because the television content is distributed to each subscriber over (conceptually) a single extremely long electrical conductor. The wire carries many different television programs by sending each of them at a different frequency, and a tuner at the home set-top box selects one of those channels for display. This is similar to the way a television receiver tunes broadcast signals from the air (although broadcast TV uses a different set of frequencies).

39. Since at least the 1970s, cable television systems have enabled two-way data communication between a central headend and users' set-top boxes. In order to provide any form of interactive cable system some upstream channel (or "back channel") must be provided. The earliest use of such a back channel was to allow subscribers to pay to view selected broadcasts, typically sports events like

boxing matches. Later forms of interaction evolved to include video-on-demand, information retrieval, and games.

40. Early systems utilized telephone lines to send data from the user's home to the central headend. An IBM Technical Disclosure Bulletin from 1995, entitled "Speech Recognition Methods for Controlling Cable Television," illustrates remote speech recognition processing and control in this type of system utilizing a telephone line:



IBM Bulletin at Fig. 1 (Ex. 1008). In this figure, a user's house (1) has a telephone (5) that is connected to a conventional telephone network (6). IBM Bulletin at 285. The telephone company switching system (7) is connected to a speech recognition system (8), which can effectuate modifications of the television signal to the user via signal generator (9) based on the recognized speech commands. IBM Bulletin at 285.

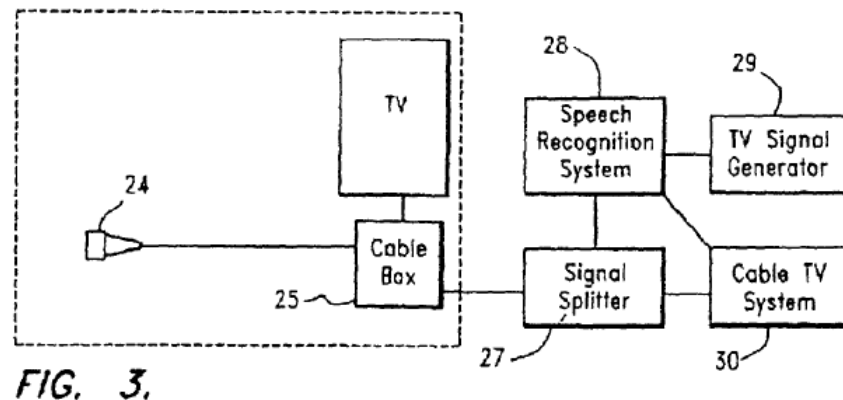
41. Other systems were soon developed using the cable television lines themselves to transmit data upstream from customer premises. One of the more noteworthy deployments was Warner's QUBE system deployed in 1977 in Columbus, Ohio. QUBE allowed a 125 kilobit/second upstream channel that permitted a wide range of interactive television applications. By 1983, Cox offered its version of two-way cable in a system called INDAX. The INDAX system used carrier sensing and collision detection for multiplexed access by its set-top boxes.

42. As interactive cable systems evolved, many used a form of time division multiple access (TDMA) multiplexing. In the TDMA approach, part of the frequency band in the cable plant was reserved for the upstream, or "back channel," as opposed to carrying video programs downstream (i.e., on the "forward channel"). The back channel was divided into many small "time slots" with each set-top box assigned a particular slot during which it was allowed to send data. This helped avoid interference with data being sent by another set-top box. Many set-top boxes would send relatively small amounts of data upstream (compared to the bandwidth of a video program sent downstream), and all the set-top boxes would be constrained to share the same upstream frequencies.

43. In the mid-1990s, cable set-top boxes were not powerful enough for computationally intensive speech recognition, as they faced hardware constraints

similar to phones and PCs. Accordingly, the cable television industry explored server-based voice recognition approaches utilizing the existing two-way data communication paths that enabled interactive television services.

44. The same 1995 IBM Technical Disclosure Bulletin discussed above also described this approach of employing a remote speech recognition system in a cable television network utilizing the two-way communication available in such a network:



IBM Bulletin at Fig. 3; see also 286-287 (Ex. 1008). In this figure, a microphone (24) connected to the cable box (25) receives voice commands. IBM Bulletin at 287. The microphone could be a wired or wireless telephone handset connected to the cable box (25). IBM Bulletin at 287. From the cable box, acoustic data from the user is sent over the cable TV network, either compressed or intermixed with transmissions from other users. IBM Bulletin at 287. Voice commands are directed to a speech recognition system (28), which controls the generation of the

television signal. IBM Bulletin at 287.

45. By the 1990s, it was apparent that in cable television systems, processing-intensive speech recognition would need to be performed at a centralized location on a powerful computer rather than on local set-top boxes with relatively low processing power. For example, engineers working at OpenTV filed a patent application in 1998 making this exact point:

While [speech and handwriting recognition] technologies have evolved to the point that they can reliably generate accurate textual information from the user's voice or handwriting, they require a great deal of computer resources. The applications may be very large and they need large amounts of processing time to perform the required pattern matching. These technologies therefore cannot be implemented in current set-top boxes, which have very limited storage and processing capacity.

Dureau at 2:36-44 (Ex. 1009). Accordingly, the OpenTV applicants described a cable television system wherein voice commands, received by a remote control with a microphone, would be “transmitted to the server, which uses voice recognition software to convert the voice data into textual data.” Dureau at 10:56-67. The recognized speech, now in textual form, could then either be sent back to the set-top box or used at the server. Dureau at 3:9-11, 3:25-27, 10:67-11:1.

46. Thus, by the end of the 1990s, server-based speech recognition systems were well known, and cable television systems in particular had been

implemented with voice-interfaces. Further, the notion of locating the speech recognition processing at a central location, such as the cable headend, was well known.

V. THE '196 PATENT

A. Overview of the '196 Patent

47. The '196 Patent is entitled “System and Method of Voice Recognition Near a Wireline Node of a Network Supporting Cable Television and/or Video Delivery.” '196 Patent at cover (Ex. 1001). The patent discloses a “method and system of speech recognition presented by a back channel from multiple user sites within a network supporting cable television and/or video delivery” '196 Patent at Abstract; see also 5:3-5 (“An embodiment of the invention provides speech recognition services to a collection of users over a network that supports cable television and/or video delivery.”).

48. The '196 Patent states that prior art “speech operated functions using the latest voice recognition technologies are limited to a handful of applications” running on systems that “typically offer only a limited number of commands” with “recognition efficiency [that] is only fair and often require[s] speech training.” '196 Patent at 1:31-40 (Ex. 1001).

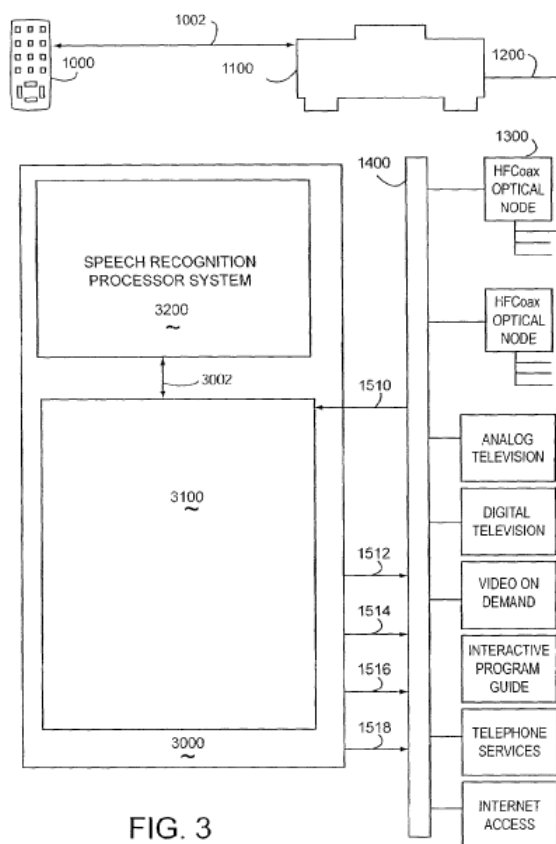
49. The '196 Patent also identifies prior art “natural language” speech recognition technology that required “state of the art processing software and

hundreds of megabytes of RAM.” ’196 Patent at 1:53-56 (Ex. 1001). The patent states that natural language recognition was already used in certain “high end systems,” but asserts that “the problems of speech recognition at a centralized wireline node in a network supporting video delivery or cable television delivery have not been addressed by such prior art.” ’196 Patent at 1:56-60, 1:63-66. According to the ’196 Patent, “[t]here is no present system providing speech recognition to a collection of users over a cable television network.” ’196 Patent at 4:59-60.

50. The ’196 Patent discloses a system that “comprises a multi-user control system for audio visual devices that incorporates a speech recognition system that is centrally located in or near a wireline node, and which may include a Cable Television (CATV) Headend.” ’196 Patent at 5:11-15 (Ex. 1001). The patent asserts that the disclosed “system is unique in that the speech command which originates at the user site, often the home of the subscriber, is sent upstream via the return path (often five to 40 MHz) in the cable system to a central speech recognition and identification engine.” ’196 Patent at 5:22-26. This engine “is capable of processing thousands of speech commands simultaneously and offering a low latency entertainment, information, and shopping experience to the user or subscriber.” ’196 Patent at 5:26-31. In the disclosed system, “spoken commands from a cable subscriber are recognized and then acted upon to control the delivery

of entertainment and information services, such as Video On Demand, Pay Per View, Channel control, on-line shopping, and the Internet.” ’196 Patent at 5:18-22.

51. Figure 3 of the ’196 Patent illustrates an embodiment “in accordance [with] the invention.” ’196 Patent at 7:25, 7:29-31 (Ex. 1001).



’196 Patent at Fig. 3.

52. As illustrated in Figure 3, “remote control unit 1000” is coupled to “set-top apparatus 1100.” ’196 Patent at 9:48-49 (Ex. 1001). “Set-top apparatus 1100 communicates through distributor node 1300 across a high-speed physical transport 1400 to a tightly coupled server farm 3000, possessing various delivery

points 1510 and entry points 1512-1518.” ’196 Patent at 9:52-55. The system includes a central “speech recognition processor system 3200.” ’196 Patent at Fig. 3. The speech processor system “may be centrally located in or near a wireline node, which may include a Cable Television (CATV) central location,” or it “may be centrally located in or near a server farm,” “web-site hosting locations,” or “gateway.” ’196 Patent at 18:24-27.

53. Remote control unit 1000 can include a microphone for receiving voice commands. ’196 Patent at 10:22-23 (Ex. 1001). The remote “relays the subscriber’s speech commands to the central speech recognition engine.” ’196 Patent at 10:26-27. The “[s]peech commands from the subscriber may be preprocessed” so that “analog signals picked up from the microphone are converted to digital signals where they undergo additional processing before being transmitted to the speech recognition and identification engine located in the cable Headend or other centralized location.” ’196 Patent at 10:54-60.

B. Prosecution History of the ’196 Patent

54. The ’196 Patent issued from U.S. Patent Application No. 09/785,375, which was filed on February 16, 2001. File History at 1-187 (Ex. 1002). The ’196 Patent claims priority to Provisional Application No. 60/210,440, filed on June 8, 2000. File History at 5; ’196 Patent at cover (Ex. 1001). The application for the ’196 Patent initially contained 66 claims. File History at 117-137.

55. On March 28, 2005, the examiner indicated that the application was in condition for allowance subject to certain objections. File History at 315 (Ex. 1002). The examiner objected to the Abstract. File History at 316. The examiner also objected to the disclosure and all claims because “the term ‘voice recognition’ is misused for what nowadays is called—**speech recognition**—in the speech signal processing art.” File History at 316 (emphasis in original). The examiner further stated: “While ‘voice recognition’ and ‘speech recognition’ were both once used interchangeably to refer to spoken word recognition, nowadays the terms are distinguished. The term ‘**voice** recognition’ now denotes identification of **who** is doing the speaking (class 704/246), while ‘**speech** recognition’ (or ‘word recognition’) denotes identification of what is being said (class 704/251).” File History at 316-317 (emphasis in original).

56. In response, on April 27, 2005, the applicants submitted a “replacement abstract” and a “replacement specification in which the term ‘voice recognition’ has been replaced with the term --speech recognition--.” File History at 332, 340-429 (Ex. 1002). The applicants confirmed using the “term ‘speech recognition’ to denote identification of what is being said.” File History at 332. The applications also noted “two instances in the specification where the Applicant refers to ‘voice print’ technology ... referring to identification of who is doing the speaking,” but clarified that “these two instances are not concerned with the

subject matter set forth in the independent claims.” File History at 332.

57. On September 7, 2005, the examiner issued a Notice of Allowance. File History at 432-435 (Ex. 1002). The patent issued on May 16, 2006. ’196 Patent at cover (Ex. 1001). On September 5, 2017, the Patent and Trademark Office issued a Certificate of Correction making a number of requested modifications to the drawings, specification, and claims. File History at 469-470.

C. Claims of the ’196 Patent

58. The ’196 Patent contains 66 claims, including 4 independent claims (i.e., 1, 14, 27, and 53).

59. There is overlap among the different claims. For instance, independent claims 1 and 27 recite several similar elements. The following chart compares claim 1 and claim 27 with differences in bold and underlined:

Claim 1	Claim 27
A <u>method of using a back channel containing a multiplicity of identified speech channels from a multiplicity of user sites presented to a speech processing system at a wireline node in a network supporting at least one of cable television delivery and video delivery</u> , comprising <u>the steps of</u> :	A <u>system supporting speech recognition in a network, said system</u> comprising:

Claim 1	Claim 27
	<u>a speech recognition system coupled to a wireline node in said network for receiving a back channel from a multiplicity of user sites coupled to said network, further comprising</u>
receiving said back channel to create a received back channel;	<u>a back channel receiver, for</u> receiving said back channel to create a received back channel;
partitioning said received back channel into a multiplicity of received identified speech channels;	<u>a speech channel partitioner, for</u> partitioning said received back channel into a multiplicity of received identified speech channels; and
	<u>a processor network executing a program system comprised of program steps residing in memory accessibly coupled to at least one computer in said processor network;</u>
processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content; <u>and</u>	<u>wherein said program system is comprised of the program steps of:</u> processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech

Claim 1	Claim 27
	content;
responding to said identified speech content to create an identified speech content response <u>that is unique</u> , for each of said multiplicity of identified speech contents.	responding to said identified speech content to create an identified speech content response, for each of said multiplicity of <u>said</u> identified speech contents; <u>and</u>
	<u>wherein said network supports at least one of the collection comprising: cable television delivery to said multiplicity of user sites; and video delivery to said multiplicity of user sites.</u>

As indicated in the chart, claim 27 recites a “system” that performs the steps recited in the “method” of claim 1.

60. Several of the dependent claims are also quite similar. The following chart compares dependent claims 2 and 28 with differences in bold and underlined:

Claim 2	Claim 28
The <u>method</u> of claim <u>1</u> ,	The <u>system</u> of claim <u>27</u> ,
wherein at least one of said identified speech channels has an associated user site; and	wherein at least one of said identified speech channels has an associated user site; and

Claim 2	Claim 28
<p>wherein <u>the step</u> partitioning said received back channel is further comprised of <u>the step</u> of:</p> <p>partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site;</p>	<p>wherein partitioning said received back channel is further comprised of:</p> <p>partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site;</p>
<p>wherein the step processing said multiplicity of said received identified speech channels is further comprised [of] the step of:</p> <p>processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech contents; and</p>	<p>wherein the <u>program</u> step processing said multiplicity of said received identified speech channels is further comprised of the <u>program</u> step of:</p> <p>processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech content; and</p>
<p>wherein the step responding to said identified speech content is further comprised [of] the step of:</p> <p>responding to said identified speech content from said associated user site to create said identified speech content</p>	<p>wherein the <u>program</u> step responding to said identified speech content is further comprised of the <u>program</u> step of:</p> <p>responding to said identified speech content from said associated user site to create said identified speech content</p>

Claim 2	Claim 28
response for said associated user site.	response for said associated user site.

As indicated in the chart, claims 2 and 28 add the same limitations.

61. The following chart compares dependent claims 4 and 30 with differences in bold and underlined:

Claim 4	Claim 30
The <u>method</u> of claim <u>2</u> , wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:	The <u>system</u> of claim <u>28</u> , wherein the <u>program</u> step responding to said identified speech contents from said associated user site is comprised of the <u>program</u> steps of:
processing said identified speech content response to create said identified user site response; and	processing said identified speech content response to create said identified user site response; and
sending said identified user site response to said identified user site.	sending said identified user site response to said identified user site.

As indicated in the chart, claims 4 and 30 add the same limitations.

62. The following chart compares dependent claims 5 and 31 with differences in bold and underlined:

Claim 5	Claim 31
The <u>method</u> of claim <u>2</u> , wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:	The <u>system</u> of claim <u>28</u> , wherein the <u>program</u> step responding to said identified speech contents from said associated user site is comprised of the <u>program</u> steps of:
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and	assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and
billing said user site based upon said financial consequence.	billing said user site based upon said financial consequence.

As indicated in the chart, claims 5 and 31 add the same limitations.

63. The following chart compares dependent claims 6 and 32 with differences in bold and underlined:

Claim 6	Claim 32
The <u>method</u> of claim <u>2</u> , wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:	The <u>system</u> of claim <u>28</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised of the <u>program</u> steps of:

Claim 6	Claim 32
assessing said speech response to create a financial consequence identified with said user site;	assessing said speech response to create a financial consequence identified as to said user site;
displaying said financial consequence to create a displayed financial consequence at said user site;	displaying said financial consequence to create a displayed financial consequence at said user site;
confirming said displayed financial consequence from said user site to create a financial commitment; and	confirming said displayed financial consequence from said user site to create a financial commitment; and
billing said user site based upon said financial commitment.	billing said user site based upon said financial commitment.

As indicated in the chart, claims 6 and 32 add the same limitations.

64. The following chart compares dependent claims 12 and 38 with differences in bold and underlined:

Claim 12	Claim 38
The <u>method</u> of claim <u>2</u> , wherein the step responding to said identified speech content from said associated user site is <u>further</u> comprised of the step of:	The <u>system</u> of claim <u>28</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised, <u>for at least one of said user sites</u> , of the

Claim 12	Claim 38
	<u>program</u> step of:
responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech content identified with said user site.	responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech contents identified as to said user site.

As indicated in the chart, claims 12 and 38 add the same limitations.

65. The following chart compares dependent claims 13 and 39 with differences in bold and underlined:

Claim 13	Claim 39
The method of claim <u>2</u> , wherein the step processing said multiplicity of said received speech channels is further comprised for at least one of said user sites is further comprised of the step of:	The <u>system</u> of claim <u>28</u> , wherein the <u>program</u> step processing said multiplicity of said received speech channels is further comprised, for at least one of said user sites, of the <u>program</u> step of:
processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.	processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.

As indicated in the chart, claims 13 and 39 add the same limitations.

66. There is substantial overlap between independent claims 14 and 53 as shown in the following chart:

Claim 14	Claim 53
A <u>program system controlling</u> at least part of a speech recognition system coupled to a wireline node in a network, <u>said program system</u> comprising the <u>program</u> steps of:	A <u>method of operating</u> at least part of a speech recognition system coupled to a wireline node in a network, comprising the steps of:
processing a multiplicity of received identified speech channels to create a multiplicity of identified speech content; and	processing a multiplicity of received identified speech channels to create a multiplicity of identified speech content; and
responding to said identified speech content to create an identified speech content response that is unique to each of said multiplicity of identified speech contents;	responding to said identified speech content to create an identified speech content response that is unique to each of said multiplicity of identified speech contents;
wherein said speech recognition system is provided said multiplicity of received identified speech channels based upon a received back channel at said wireline node from a multiplicity of user sites	wherein said speech recognition system is provided said multiplicity of received identified speech channels based upon a received back channel at said wireline node from a multiplicity of user sites

Claim 14	Claim 53
coupled to said network;	coupled to said network;
<u>wherein each of said program steps reside in memory accessibly coupled to at least one computer included in said speech recognition system;</u>	
<u>wherein said at least one computer communicatively couples through said wireline node to said multiplicity of user sites; and</u>	
wherein said network supports at least one of the collection comprising: cable television delivery to said multiplicity of user sites; and video delivery to said multiplicity of user sites.	wherein said network supports at least one of the collection comprising: cable television delivery to said multiplicity of user sites; and video delivery to said multiplicity of user sites.

As indicated in the chart, claims 14 and 53 recite essentially the same elements except that claim 53 is a method claim while claim 14 recites a “program system” that resides on a computer connected to a “multiplicity of user sites” and performs the same recited steps.

67. The dependent claims that depend from claims 14 and 53 recite essentially the same limitations. The following chart compares dependent claims 15 and 54 with differences in bold and underlined:

Claim 15	Claim 54
The <u>program system</u> of claim 14 ,	The <u>method</u> of claim 53 ,
wherein at least one of said identified speech channels has an associated user site; and	wherein at least one of said identified speech channels has an associated user site; and
wherein the <u>program</u> step processing said multiplicity of said received identified speech channels is further comprised of the <u>program</u> step of:	wherein the step processing said multiplicity of said received identified speech channels is further comprised of the step of:
processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech content; and	processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech content; and
wherein the <u>program</u> step responding to said identified speech content is further comprised of the <u>program</u> step of:	wherein the step responding to said identified speech content is further comprised of the step of:
responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.	responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.

As indicated in the chart, claims 15 and 54 add the same limitations.

68. The following chart compares dependent claims 17 and 64 with differences in bold and underlined:

Claim 17	Claim 64
The <u>program system</u> of claim <u>15</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised of the <u>program</u> steps of:	The <u>method</u> of claim <u>54</u> , wherein the step responding to said identified speech content from said associated user site is comprised of the steps of:
processing said identified speech content response to create said identified user site response; and	processing said identified speech content response to create said identified user site response; and
sending said identified user site response to said identified user site.	sending said identified user site response to said identified user site.

As indicated in the chart, claims 17 and 64 add essentially the same limitations.

69. The following chart compares dependent claims 18 and 65 with differences in bold and underlined:

Claim 18	Claim 65
The <u>program system</u> of claim <u>15</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised of the	The <u>method</u> of claim <u>54</u> , wherein the step responding to said identified speech content from said associated user site is comprised of the steps of:

Claim 18	Claim 65
<u>program</u> steps of:	
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and	assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and
billing said user site based upon said financial consequence.	billing said user site based upon said financial consequence.

As indicated in the chart, claims 18 and 65 add essentially the same limitations.

70. The following chart compares dependent claims 19 and 55 with differences in bold and underlined:

Claim 19	Claim 55
The <u>program system</u> of claim <u>15</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised of the <u>program</u> steps of:	The <u>method</u> of claim <u>54</u> , wherein the step responding to said identified speech content from said associated user site is comprised of the steps of:
assessing said speech response to create a financial consequence identified as to said user site;	assessing said speech response to create a financial consequence identified as to said user site;
displaying said financial consequence to	displaying said financial consequence to

Claim 19	Claim 55
create a displayed financial consequence at said user site;	create a displayed financial consequence at said user site;
confirming said displayed financial consequence from said user site to create a financial commitment; and	confirming said displayed financial consequence from said user site to create a financial commitment; and
billing said user site based upon said financial commitment.	billing said user site based upon said financial commitment.

As indicated in the chart, claims 19 and 55 add essentially the same limitations.

71. The following chart compares dependent claims 25 and 61 with differences in bold and underlined:

Claim 25	Claim 61
The <u>program system</u> of claim <u>15</u> , wherein the <u>program</u> step responding to said identified speech content from said associated user site is comprised, for at least one of said user sites, of the <u>program</u> step of:	The <u>method</u> of claim <u>54</u> , wherein the step responding to said identified speech content from said associated user site is comprised, for at least one of said user sites, of the step of:
responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech contents	responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech contents

Claim 25	Claim 61
identified as to said user site.	identified as to said user site.

As indicated in the chart, claims 25 and 61 add essentially the same limitations.

72. The following chart compares dependent claims 26 and 62 with differences in bold and underlined:

Claim 26	Claim 62
The <u>program system</u> of claim <u>15</u> , wherein the <u>program</u> step processing said multiplicity of said received speech channels is further comprised, for at least one of said user sites, of the <u>program</u> step of:	The <u>method</u> of claim <u>54</u> , wherein the step processing said multiplicity of said received speech channels is further comprised, for at least one of said user sites, of the step of:
processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.	processing said received speech channel from said user site based upon natural language for said user site to create said speech content identified as to said user site.

As indicated in the chart, claims 26 and 62 add essentially the same limitations.

D. Effective Filing Date of the '196 Patent

73. The '196 Patent was filed on February 16, 2001, and issued on May 16, 2006. '196 Patent at cover (Ex. 1001). The patent claims priority to

Provisional Application No. 60/210,440 (Ex. 1003) filed on June 8, 2000.

74. I have used June 8, 2000, as the effective filing date for my analysis of the asserted claims of the '196 Patent because the patent claims priority to the provisional application filed on that date. However, the provisional application does not discuss all of the elements of the challenged claims. For instance, claim 1 recites the step of "partitioning said received back channel into a multiplicity of received identified speech channels." '196 Patent at 51:3-4 (Ex. 1001). The provisional application does not discuss "partitioning" or otherwise identifying a "multiplicity of received identified speech channels." Provisional Appl. at 4-30 (Ex. 1003). Similarly, the provisional application does not discuss "processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content," which is recited by each of the challenged independent claims (i.e., claims 1, 14, and 53). '196 Patent at 51:4-6, 53:1-3, 58:15-17; Provisional Appl. at 4-30. Nor does the provisional application discuss "responding to said identified speech content to create an identified speech content response that is unique, for each of said multiplicity of identified speech contents," which is also recited by each of the challenged independent claims. '196 Patent at 51:7-9; Provisional Appl. at 4-30.

E. Level of Ordinary Skill in the Art

75. In my opinion, a person of ordinary skill in the art with regard to the

'196 Patent would have been familiar with interactive multi-media network architectures and speech recognition and voice control technologies. That familiarity would have come through having either (i) an undergraduate degree (or equivalent) in electrical engineering, computer science, or a comparable subject and at least 3 years of professional work experience in the field of multi-media systems including in particular speech recognition and control technologies, or (ii) an advanced degree (or equivalent) in electrical engineering, computer science, or a comparable subject and at least 1 year of post-graduate research or work experience in the field of multi-media systems including in particular speech recognition and control technologies.

76. The '196 Patent and the prior art references discussed in this declaration all involve interactive multi-media networks and speech recognition and voice control technologies. I was working in the interactive multi-media field in the period leading up to the filing of the '196 Patent (i.e., prior to 2000). My own work with speech recognition began in 1979 when I built the first conversational speech system (called "Put That There"), which also included multi-modal input and multi-media output. This system used the first commercially available connected speech recognizer (i.e. one that does not require a pause between each word). Subsequently through the 1980s and 1990s, I built interactive speech systems to access email and voice mail, edit dictation, control

computer window systems, power facial animations, manage shopping lists, provide access to broadcast newscasts, and to access personal calendars and address books. People working on interactive multi-media systems during the time period from roughly the early 1980s to 2000 generally had an undergraduate or graduate degree in electrical engineering, computer science, or a comparable area. My personal experience informs my opinion as to the ordinary skill in the art at the relevant time.

77. In 1996, I helped found InTouch Systems, a startup providing voice access over the telephone to subscribers to manage their voice mail and voice dial, among other features. This required precisely the sort of “natural language” speech recognition referenced in the ’196 Patent. The system vocabulary was large and included surnames, and it was not practical to train on each subscriber’s voice in advance as in many earlier systems. In this work we evaluated well-known large vocabulary speech recognition products, including BBN’s HARK, Applied Language Technology, and Nuance. These are exactly the sorts of recognizers cited in the background section of the ’196 Patent specification and were commercial products that would have been known to anyone of ordinary skill in the art of designing speech recognition and control systems for large audiences of users.

VI. CONSTRUCTION OF TERMS USED IN THE '196 PATENT

78. I understand that patent claims in an *inter partes* review are given the broadest reasonable interpretation in light of the patent specification. I further understand that claim terms are generally given their ordinary and customary meaning, which is the meaning that terms would have to a person of ordinary skill in the relevant art at the time of the claimed invention, consistent with the broadest reasonable construction.

79. I further understand that the construction of claims in patent litigation in a United States District Court would be subject to a different standard than the broadest reasonable interpretation standard that applies in an *inter partes* review. I also understand that this difference in the standards applied in court and in an *inter partes* review can lead to different constructions of the same terms. I have not attempted to determine the appropriate construction of any terms of the '196 Patent under the standard that applies in a court proceeding.

A. “wireline node” (claims 1, 14, 27, 53)

80. The challenged independent claims each refer to a “wireline node.” For instance, claim 1 recites “a speech processing system at a wireline node in a network” supporting cable television or video delivery. '196 Patent at 50:64-66 (Ex. 1001). Similarly, claims 14 and 53 recite “a speech recognition system coupled to a wireline node in a network.” '196 Patent at 52:65-66, 58:12-13. The

patent specification states that “a centralized wireline node refers to a network node providing video or cable television delivery to multiple users using a wireline physical transport between those users at the node.” ’196 Patent at 1:66-2:2. In a cable television system, this could be the headend serving multiple subscribers over coaxial and fiber-optic cables.

81. In my opinion, based on the patent specification and general knowledge in the field, a person of ordinary skill in the art would understand the term “wireline node” as used in the challenged claims to mean “a network node providing video or cable television delivery to multiple users over physical wires between the node and the multiple user locations.”

B. “back channel” (claims 1, 2, 14, 27, 28, 53)

82. The challenged claims recite a “back channel.” For instance, claim 1 recites the steps of “receiving said back channel to create a received back channel” and “partitioning said received back channel into a multiplicity of received identified speech channels.” ’196 Patent at 51:1-4 (Ex. 1001). As I discussed above, cable television systems have for decades enabled two-way data communication between a central headend and users’ set-top boxes with an upstream channel (or “back channel”) used for signal transmission from the user to the headend. See Paragraphs 39-44, above. The ’196 Patent uses the term in this generally understood way. For example, the patent states that “[t]he back channel

is from a multiplicity of user sites and is presented to a speech processing system at the wireline node in the network.” ’196 Patent at 22:12-14. The specification also states, in reference to Figures 21 and 22, that “Uplink 1204 communication includes a back channel” and that “This back channel includes multiple identified speech channels from multiple user sites (STBs) 1100” ’196 Patent at 30:4-7; see also 32:37-39 (“Uplink 138 communication includes a back channel. This back channel includes multiple response channels from multiple user sites STBs, as shown in FIGS. 1 and 2.”).

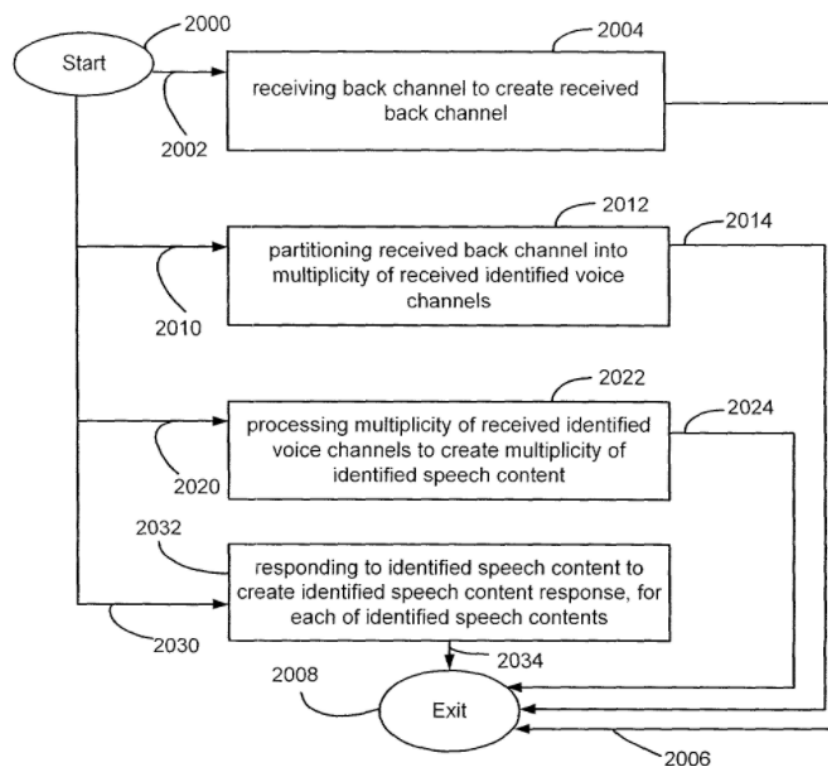
83. In my opinion, based on the patent specification and general knowledge in the field, a person of ordinary skill in the art would understand the term “back channel” to refer to the “upstream communication channel delivering signals from multiple user sites to a central wireline node.”

C. “partitioning said received back channel into a multiplicity of [said] received identified speech channels” (claims 1, 2, 27, 28)

84. Claims 1 and 27 recite the step of “partitioning said received back channel into a multiplicity of received identified speech channels.” ’196 Patent at 51:3-4 (Ex. 1001). Dependent claims 2 and 28 further limit this step by reciting “partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site.” ’196 Patent at 51:14-18 (emphasis added).

85. Neither the claims nor the patent specification describe any particular

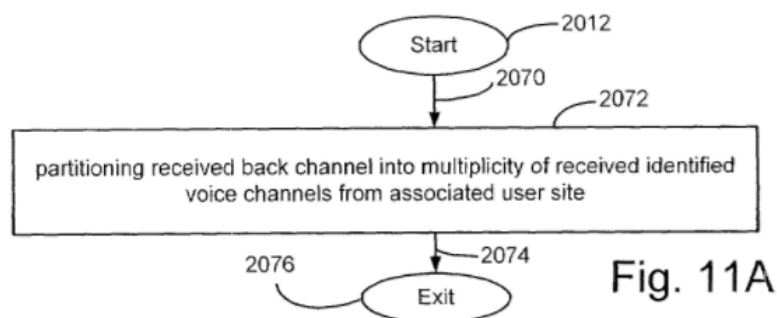
approach to “partitioning” the back channel. The specification simply restates the same language used in the claims without further explanation. For example, Figure 10 “depicts a flowchart of a method that uses a back channel containing a multiplicity of identified speech channels for speech recognition at a wireline node in a network supporting video or cable television delivery in accordance with the invention.” ’196 Patent at 22:8-12 (Ex. 1001).



’196 Patent at Fig. 10. Step 2012 in the figure recites the “partitioning” step of claim 1 in nearly identical terms. ’196 Patent at 51:3-4 (“partitioning said received back channel into a multiplicity of received identified speech channels”). The associated text in the specification provides no further explanation, stating only

that “[o]peration 2012 performs partitioning the received back channel into a multiplicity of received identified speech channels.” ’196 Patent at 22:24-26.

86. Figure 11A “depicts a detail flowchart of operation 2012” shown in Figure 10. ’196 Patent at 22:66-67 (Ex. 1001). But Figure 11A simply recites the same language included in claim 2.

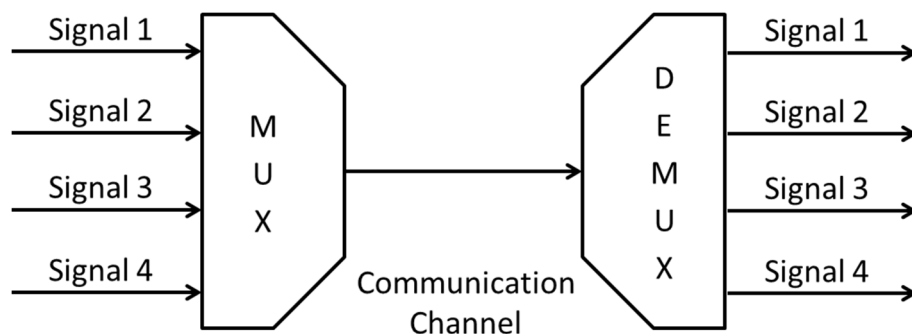


’196 Patent at Fig. 11A. And again the associated text in the specification adds no further detail, stating that “[o]peration 2072 performs partitioning the received back channel into a multiplicity of received identified speech channels from the associated user site.” ’196 Patent at 23:2-5.

87. Because neither the claims nor the specification specifically describe or limit the term “partitioning” to any particular approach or technique, it is my opinion that a person of ordinary skill in the art would understand the phrase “partitioning said received back channel” broadly to include any technique used to divide a communication channel into identifiable portions. Thus, in the context of claims 1 and 2, a person of ordinary skill in the art would understand the phrase “partitioning said received back channel into a multiplicity of [said] received

identified speech channels” to mean “dividing the received back channel into multiple portions that are each identified with speech commands originating from system users”.

88. In my opinion, one well-known technique for “partitioning” a communication channel – as the term is used in the ’196 Patent – would involving multiplexing (i.e., combining) several independent communication signals together for transmission to the central network node and then de-multiplexing (i.e., decomposing) that signal into portions identified with each sending user location. The ’196 Patent specification discusses multiplexing in the “Background of the Invention” section, and I agree that it was a well-known technique used in many different types of communication systems long before the ’196 Patent. ’196 Patent at 3:44-52 (Ex. 1001). The following diagram illustrates the basic concept:



The multiplexer (MUX) combines multiple signals for transmission on the communication channel, and the demultiplexer (DEMUX) disaggregates the signal received on the communication channel to recreate the multiple input signals. This

technique effectively partitions the communication channel so that the receiving device can identify and appropriately process the different signals multiplexed together by the transmitting device.

89. In my opinion, this technique of de-multiplexing a multiplexed signal is one example of “partitioning” as that term is used in the challenged claims. It would have been well-known to a person of ordinary skill in the art long before the ’196 Patent.

VII. OVERVIEW OF PRIOR ART REFERENCES

A. Overview of Murdock

90. U.S. Patent No. 7,013,283, which issued to Murdock and others, is entitled “System and Method for Providing Programming Content in Response to an Audio Signal.” Murdock at cover (Ex. 1010). Murdock was filed on November 16, 2000, and issued on March 14, 2006. It claims priority to Provisional Application No. 60/166,010, which was filed November 17, 1999. I understand that Murdock is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

91. Murdock discloses “a system and method for providing video content in response to an audio signal such as a spoken audio command.” Murdock at 1:41-43 (Ex. 1010). In the Murdock invention, “programming content and the audio signal are transmitted in a network having a forward channel and a back

channel.” Murdock at 1:43-45. A “local processing unit” receives the first audio signal and transmits the received first audio signal via the back channel to a remote server computer. Murdock at 1:49-50. The remote server “receives the first audio signal from the back channel” and then “recognizes the first user and the request for programming content.” Murdock at 1:51-53. It then “retrieves the requested programming content from a program database and transmits the programming content from the local processing unit via the forward channel.” Murdock at 1:53-56.

92. Figure 1 shows a block diagram of the “voice activated system 100” of the present invention. Murdock at 2:11-12 (Ex. 1010).

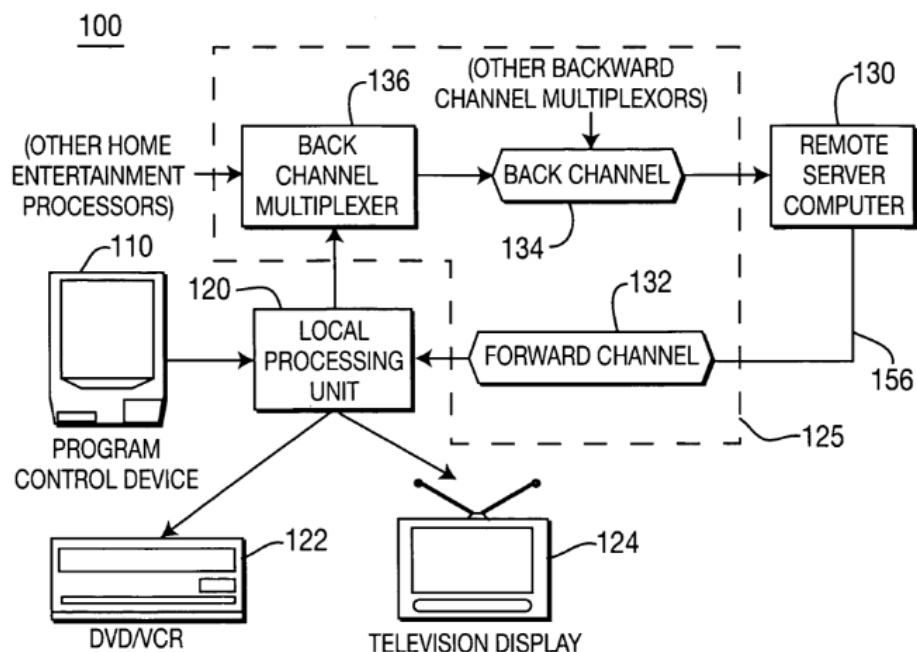


FIG. 1

Murdock at Fig. 1. In one embodiment, “the program control device 110

comprises a portable or hand-held controller.” Murdock at 2:35-36. The program control device 110 “captures the input verbal command from the user of the voice activated control system 100.” Murdock at 2:22-24. This verbal command “may comprise a verbal request of programming content from a service provider.”

Murdock at 2:24-26. After receiving the command, the program control device 110 “performs a transmission, e.g., a wireless transmission, of the command signal to the local processing unit 120.” Murdock at 2:31-34.

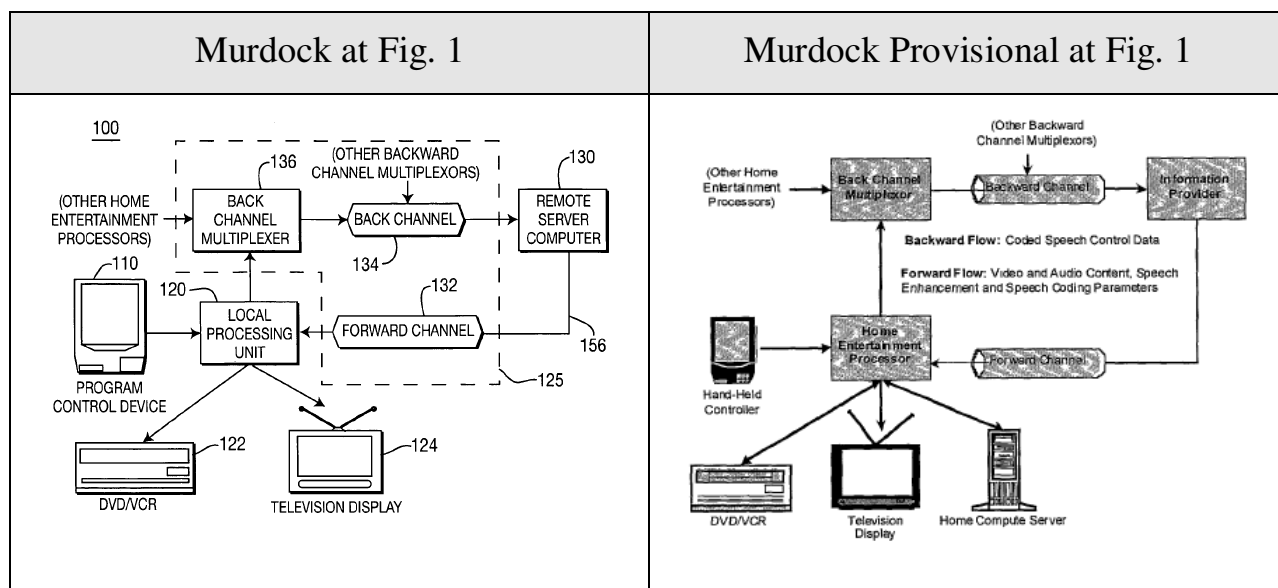
93. The “local processing unit 120,” which may include a “set top terminal, a cable box, and the like,” receives the input command from the program control device 110.” Murdock at 2:44-47 (Ex. 1010). The processing unit 120 identifies the user using either audio and/or video data. Murdock at 2:49-51 (“In one embodiment, the local processing unit 120 identifies the user upon receipt of the audio signal.”); 2:57-61 (“[I]f the received input command comprises a video signal, the local processing unit 120 extracts visual information of the user from the video signal and identifies the user from the extracted information, e.g., lip location of the user.”). If the user is verified, the local processing unit 120 “transmits the audio signal to the back channel multiplexer 136.” Murdock at 2:51-53.

94. “The back channel multiplexer 136 multiplexes or combines the audio signal transmitted from the local processing unit 120 with audio signals from the

local processing units of other users.” Murdock at 3:1-5 (Ex. 1010). “The multiplexed signal is then transmitted via the back channel 134 to the service provider.” Murdock at 3:5-7.

95. At the service provider is a remote server computer 130. Murdock at 3:9-12 (Ex. 1010). The remote server computer 130 “receives the multiplexed signal from the back channel 134 and performs speech recognition on the received signal.” Murdock at 3:15-17. It performs the speech recognition “by identifying or recognizing the user that generated the audio signal and determining the requested programming content contained in the audio signal.” Murdock at 3:28-31. “Once the user is identified and the requested programming content is determined, the server computer 130 retrieves the requested program content from a program database and transmits retrieved program content via the forward channel 132 to the local processing unit 120.” Murdock at 3:31-36.

96. As I stated above, Murdock is based on a provisional patent application filed on November 17, 1999. Murdock Provisional (Ex. 1011). The Murdock Provisional discloses the same system disclosed and claimed in the Murdock patent, as illustrated by the following figures:



97. Both Murdock and the Murdock Provisional disclose a system permitting a user to use voice commands to obtain video programming from a remote service provider. Murdock at 1:8-11 (“The invention relates generally to a system and a concomitant method for audio processing and, more particularly, to a system and method for providing video content in response to an audio signal.”) (Ex. 1010); Murdock Provisional at 1:4-6 (“The present invention relates to an apparatus, system and concomitant method for voice control of a networked television-centric information service that utilizes a backward channel capability.”) (Ex. 1011).

98. I understand that in order for the Murdock patent to have the priority date established by the Murdock Provisional at least one claim of the Murdock patent must be supported by the disclosure in the Murdock Provisional. I have analyzed the claims of Murdock and the content of the Murdock Provisional, and

in my opinion the Murdock Provisional does support the claims of the Murdock patent. To show how I reached my conclusion, I provide the following element-by-element comparison of claim 1 in Murdock with the Murdock Provisional.

99. Claim 1 of Murdock begins with the following preamble:

A system for providing programming content in response to an audio signal, where said audio signal and programming content are transmitted using a television network having a forward channel and a back channel, the system comprising:

100. The Murdock Provisional describes a system for providing programming content in response to an audio signal, as recited in the preamble of claim 1. Murdock Provisional at 1:4-6 (“The present invention relates to an apparatus, system and concomitant method for voice control of a networked television-centric information service that utilizes a backward channel capability.”) (Ex. 1011). The audio signal and programming content are transmitted using a television network with a forward channel and back channel, as recited in the preamble of claim 1. Murdock Provisional at 1:18-22 (“The central speech recognition server is contained at the network information source. It processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.”), 2:3-7 (“The Home Entertainment Processor (HEP) is the core component of a home entertainment system which is primarily a television viewing

system. The HEP is an electronic device for receiving network programming content through the forward channel and transmitting speech control data through the backward channel.”), 4:1-3, 4:18-22. For instance, the provisional application is entitled: “A System and Method for Voice Navigation for Networked Television-Centric Information Services with Back Channel,” Murdock Provisional at 1:1-2. Additionally, Figure 1 in the provisional application illustrates a “television network” and specifically identifies a “Forward Channel” and a “Backward Channel.” Murdock Provisional at Fig. 1.

101. The first element of claim 1 of the Murdock patent reads as follows:

a local processing unit for receiving a first audio signal from a first user, where said first audio signal contains a request for said programming content from a service provider, transmitting said first audio signal to the service provider via a back channel of said television network;

102. The Murdock Provisional discloses this element. The “Home Entertainment Processor” (or HEP) in the provisional application is the “local processing unit” recited in claim 1 of the Murdock patent. As the Murdock Provisional explains: “The Home Entertainment Processor (HEP) is the core component of the home entertainment system which is primarily a television viewing system. The HEP is an electronic device for receiving network programming content through the forward channel and transmitting speech control

data through the backward channel.” Murdock Provisional at 2:3-7 (Ex. 1011).

103. The HEP receives an audio signal from a user, as recited in the claim. In particular, the disclosed system includes a “Hand Held Controller (HHC)” that is “used by the viewer in a manner that is similar to the way current television remote controls are used with the exception that the viewer makes selections by speaking into the HHC rather than pushing buttons.” Murdock Provisional at 2:16-19 (Ex. 1011). “The viewer’s speech command (‘Show me all of the NCAA football games on this afternoon’) is acquired from the HHC with one or more microphones, which is then transmitted to the HEP.” Murdock Provisional at 2:19-22; see also 3:8-11 (Hand Held Controller “contains one or more microphones, a video camera and a radio-frequency transmitter for sending the speech and video data to the HEP.”).

104. The user’s voice signal can request programming content from a service provider. For instance, the provisional application provides the specific example of a viewer speaking “Show me all of the NCAA football games on this afternoon.” Murdock Provisional at 2:19-20 (Ex. 1011). The voice request is sent to “the information source,” which can provide “web-content, video-on-demand, cable television, or a local cassette player.” Murdock Provisional at 2:9-14. The information source “is located at the site of the information provider.” Murdock Provisional at 4:1-3.

105. The audio signal is transmitted to the service provider using the back channel of the network, as recited in claim 1. Murdock Provisional at 1:15-18 (“It is the second advantage of this invention to provide a mechanism, the network’s backward channel, for transmitting voice control data to the network information source.”), 2:5-7 (“The HEP is an electronic device for ... transmitting speech control data through the backward channel.”), 2:9-11 (“The back channel is the channel in which the Voice Control Data from the HEP is transmitted to the information source.”), 2:28-29 (“[T]he speech signal is coded for efficient transport over the back channel and sent out to the Back Channel Multiplexor.”), 3:23-26 (“The coded speech commands are sent out to the network’s backward channel in order that the speech commands can be processed by a speech recognition engine at the information source.”), 3:28-30 (Backward Channel Multiplexor “mixes the backward channel content received from several HEPs into one signal that is sent to the Information source.”) (Ex. 1011).

106. The second element of claim 1 of the Murdock patent reads as follows:

a remote server computer for receiving said first audio signal from the back channel, recognizing the first user and said request for said programming content from said transmitted audio signal, retrieving the request programming content from a program database, and transmitting said programming content to said

local processing unit via the forward channel.

107. The Murdock Provisional discloses a “remote server computer” as recited in claim 1 of the Murdock patent. In particular, the Murdock Provisional discloses an “Information Source” that “is located at the site of the information provider” and “contains the central speech recognition server.” Murdock Provisional at 4:1-3 (Ex. 1011); see also 4:18-22 (“The central speech recognition server is contained at the network information source. It processes speech control data arriving through the backward channel”). A person of ordinary skill would understand that the central speech recognition server is either itself a remote server computer or that it executes on a remote server computer or both.

108. The remote server computer receives the audio signal from the back channel, as recited in the claim. The Murdock Provisional explains that the “central speech recognition server ... processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 4:18-22 (Ex. 1011); see also 1:15-18 (“It is the second advantage of this invention to provide a mechanism, the backward channel, for transmitting voice control data to the network information source.”), 2:5-7 (“The HEP is an electronic device for ... transmitting speech control data through the backward channel.”), 2:9-11 (“The back channel is the channel in which the Voice Control

Data from the HEP is transmitted to the information source.”), 3:23-26 (“The coded speech commands are sent out to the network’s backward channel in order that the speech commands can be processed by a speech recognition engine at the information source.”).

109. The remote server recognizes the user and the request for programming content, as recited in the claim. The provisional application states that the Home Entertainment Processor “contains software modules that perform speaker identification, speech enhancement and coding and video tracking.” Murdock Provisional at 3:16-18 (Ex. 1011). “The speaker identification is used by the speech server to perform a more accurate recognition by employing a user profile that contains a statistical model of an individual’s preferences and command patterns.” Murdock Provisional at 3:20-23; see also 4:6-8 (“Also, an inference engine for processing user profile information is used by the Information Source to help guide the speech recognition engine.”). A person of ordinary skill in the art would understand from the disclosure in the Murdock Provisional that the central speech recognition server uses the identified speaker information from the HEP to associate a user profile with the identified speaker (i.e., based on the identification information, the central speech recognition server recognizes the speaker and pulls the speaker’s profile to help the speech recognition engine interpret the speaker’s commands). The Murdock Provisional thus discloses

“recognizing the first user” from the “transmitted audio signal,” as recited in the claim.

110. In addition, the central speech recognition server recognizes the user’s request for programming content. For example, the Murdock Provisional states that “central speech recognition server ... at the information source ... processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011). The information delivered by the information source “could be web content, video-on-demand, or cable television.” Murdock Provisional at 4:2-3; see also 2:12-14. Thus, the central speech recognition server “recognizes the user’s request for programming” (e.g., cable television programs, on-demand videos, etc.) so that the information source can then identify and deliver the requested programming.

111. The remote server retrieves the programming content requested by the user suggesting to a person of ordinary skill in the art that at least one source of the content is a program database, as recited in the claim. The Murdock Provisional states that the information delivered from the information source “could be web content, video-on-demand, or cable television.” Murdock Provisional at 4:1-3 (Ex. 1011). In order to transmit the requested programming content, the information source would have to retrieve the programming content from the

information provider. A person of ordinary skill in the art would have known that videos in a video-on-demand system would commonly be maintained in a database. Similarly, certain web content would commonly be maintained in a database (e.g., videos, images, etc.). Thus, because the Murdock Provisional discloses an information source that transmits information, such web content and videos in a video-on-demand system, to the viewer, the provisional indicates to a person of ordinary skill in the art that the programming content is obtained from a database.

112. The remote server transmits the programming content back to the local processing unit (i.e., Home Entertainment Processor or HEP) using the forward channel, as recited in claim 1. The Murdock Provisional explicitly states that “[t]he HEP is an electronic device for receiving network programming content through the forward channel” Murdock Provisional at 2:5-7; see also 1:18-22 (“The central speech recognition server is contained at the network information source. It processes speech control data arriving through the backward channel and the sends the appropriate programming modifications through the forward channel as requested by the viewer.”), 2:7-9 (“The forward channel is the channel in which the programming content from the information source is transmitted into the HEP.”), 3:15-26 (Home Entertainment Processor “is an electronic device for receiving network content through the network’s forward channel.”) (Ex. 1011).

113. Thus, as stated above, it is my opinion that the Murdock Provisional discloses support for every element of at least claim 1 of the Murdock Patent. Indeed, in my opinion, it provides support for other claims as well. For instance, claim 14 is a method claim that closely follows the elements recited in claim 1, and it is supported by the disclosure in the provisional application for the same reasons. In addition, the Murdock Provisional also discloses the subject matter that I rely on from the Murdock patent in comparing Murdock with the challenged claims. Thus, I have provided citations to the Murdock Provisional below where I compare the challenged claims to Murdock.

B. Overview of Julia

114. U.S. Patent No. 6,513,063, which issued to Julia and others, is entitled “Accessing Network-Based Electronic Information Through Scripted Online Interfaces Using Spoken Input.” Julia at cover (Ex. 1012). It issued from U.S. Patent Application No. 09/524,868, which was filed on March 14, 2000 as a continuation-in-part of U.S. Application No. 09/225,198 filed on January 5, 1999. The Julia patent claims priority to three provisional applications filed on March 17, 1999 (i.e., 60/124,718; 60/124,719; and 60/124,720). I understand that Julia is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

115. Julia discloses a “system, method, and article of manufacture for navigating network-based electronic data sources in response to spoken input

requests.” Julia at 2:27-30 (Ex. 1012). In Julia, “[t]he interpretation of the spoken request can be performed on a computing device locally with the user or remotely from the user.” Julia at 2:34-36. “The resulting interpretation of the request is thereupon used to automatically construct an operational navigation query to retrieve the desired information from one or more electronic network data sources, which is then transmitted to a client device of the user.” Julia at 2:36-41.

116. Figure 1a of Julia “illustrates a system providing a spoken natural language interface for network-based information navigation, in accordance with an embodiment of the present invention with server-side processing of requests.” Julia at 3:7-10 (Ex. 1012).

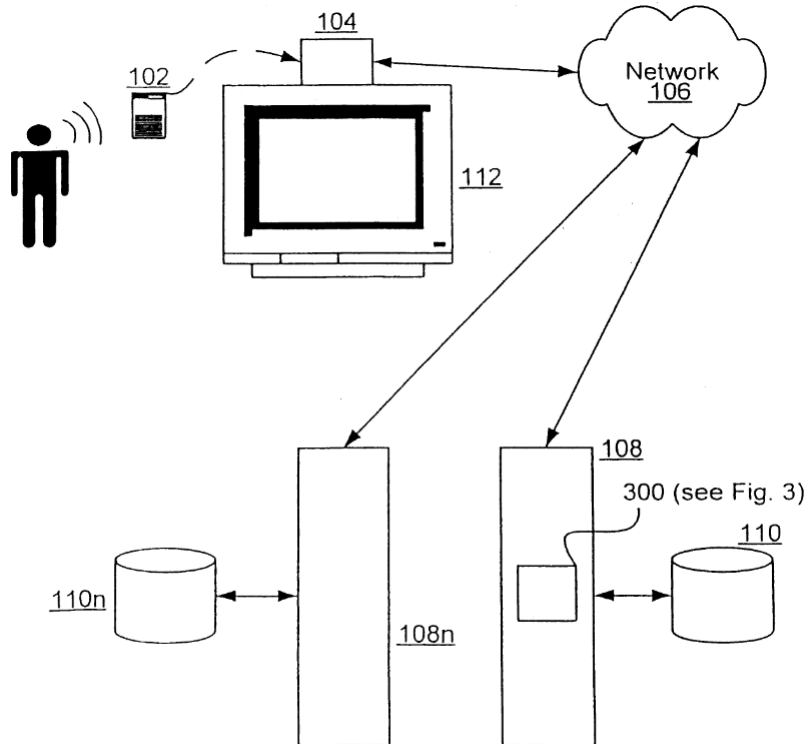


Fig. 1a

Julia at Fig. 1a. In an embodiment that is “well suited for the home entertainment setting, voice input device 102 is a portable remote control device with an integrated microphone, and the voice data is transmitted from device 102 to communication box 104” (e.g., a set-top box). Julia at 3:45-48. “The voice data is then transmitted across network 106 to a remote server or servers 108” in either digital or analog format. Julia at 3:54-60.

117. At remote server 108, “request processing logic 300” processes the voice data “in order to understand the user’s request and construct an appropriate query or request for navigation of remote data source 110.” Julia at 3:61-64 (Ex. 1012). “Data source 110 may include multimedia content, such as movies or other digital video and audio content, other various forms of entertainment data, or other electronic information.” Julia at 4:10-13. To understand the user’s request, “request processing logic 300” “comprises functional modules including speech recognition 310.” Julia at 3:66-4:2. “Speech recognition...is performed using speech recognition engine 310.” Julia at 7:12-13. A natural language interpreter (or parser) 320 “linguistically parses and interprets the textual output of the speech recognition engine.” Julia at 7:47-49. “Request processing logic 300” then “identifies and selects an appropriate online data source where the desired information ... can be found.” Julia at 8:34-37.

118. After the requested information has been retrieved, “it is electronically

transmitted via network 106 to the user for viewing on client display device 112” (e.g., “a television monitor or similar audiovisual entertainment device”). Julia at 4:18-24 (Ex. 1012).

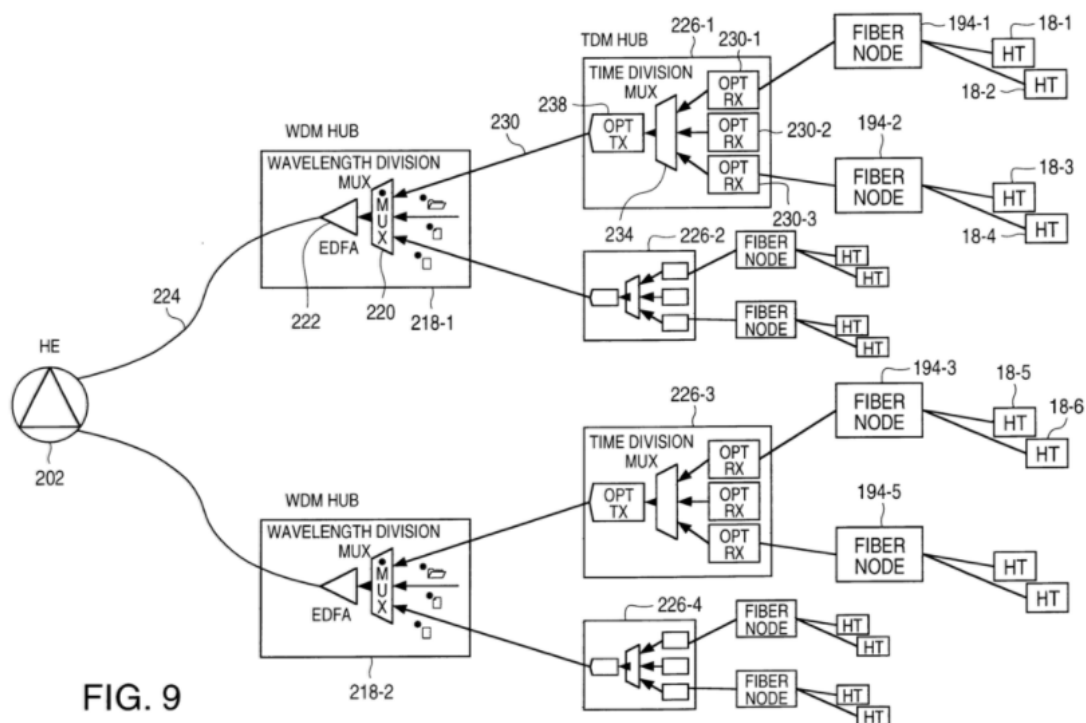
C. Overview of Nazarathy

119. U.S. Patent No. 6,490,727, which issued to Nazarathy and others, is entitled “Distributed Termination System for Two-Way Hybrid Networks.” Nazarathy at cover (Ex. 1013). It issued from U.S. Patent Application 09/414,979, which was filed on October 7, 1999. I understand that Nazarathy is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

120. According to Nazarathy, in conventional HFC systems, the conversion of analog signals from upstream communications—for example, from home set top boxes to a cable headend—typically occurred at burst receivers located within the headend. Nazarathy at 9:35-40 (Ex. 1013). In these systems, “[t]he fiber nodes are merely transparent linear points of conversion between the analog RF formats on the home side and the analog optical power formats on the head-end side. The digital detection . . . of the upstream transmission occurs in the burst receivers at the head-end or hubs.” Nazarathy at 9:41-46. Nazarathy distinguishes itself from these “conventional” systems by distributing the “functionality of the ITS head-end controller” “between the head-end and the fiber nodes, moving the upstream detection from the head-end to the fiber nodes, i.e. placing the upstream burst

receivers in the fiber nodes, transmitting the detected digital information in digital form from the fiber nodes to the head-end, and further introducing a novel distributed method to enable this by providing synchronization and calibration control between the two parts of the distributed ITS.” Nazarathy at 9:54-63.

121. Figure 9 of Nazarathy shows a hybrid TDM (Time Division Multiplexing) – WDM (Wavelength Division Multiplexing) network return path that illustrates how data from multiple home terminals 18, such as a cable modem or set-top box, is transmitted to the cable headend 202 in the invention disclosed in Nazarathy:



Nazarathy at Fig. 9 (Ex. 1013)

122. Here, RF signals from multiple home terminals (labeled “HT” in

Figure 9) propagate via coaxial cables to fiber nodes, where they are converted from RF to digital formats using burst receiver demodulators. Nazarathy at 13:63-14:13 (Ex. 1013). These signals “depart from the fiber nodes 194 in digital form on the upstream fibers, propagating all the way back to the head-end 202, possibly via intermediate levels of WDM and/or TDM multiplexing.” Nazarathy at 14:28-33. Figure 9 “depicts the combined WDM and TDM case, but elements of it could be used for the cases of WDM alone and TDM alone.” Nazarathy at 14:33-35. “Any operations of TDM and/or WDM multiplexing are undone at the HE by corresponding WDM and TDM demultiplexers,” which can be implemented by, for example, an Ethernet switch. Nazarathy at 14:62-64, 15:40-46.

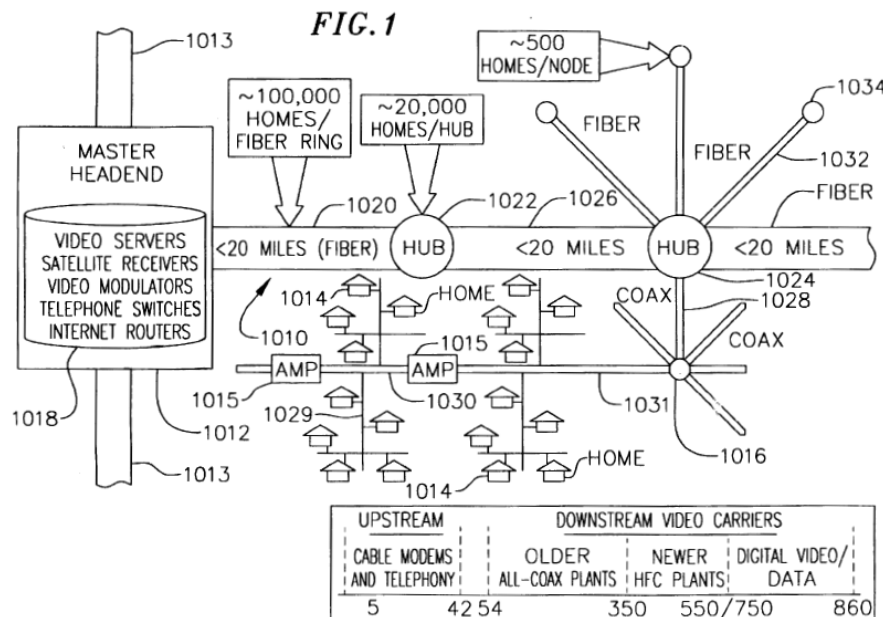
D. Overview of Quigley

123. U.S. Patent No. 6,650,624, which issued to Quigley and others, is entitled “Cable Modem Apparatus and Method.” Quigley at cover (Ex. 1014). It issued from U.S. Patent Application 09/574,558, which was filed on May 19, 2000, as a continuation of U.S. Application No. 09/430,821, which was filed on October 29, 1999. The Quigley patent claims priority to six provisional applications filed on October 30, 1998 (i.e., 60/106,264; 60/106,427; 60/106,438; 60/106,439; 60/106,440; and 60/106,441). I understand that Quigley is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

124. Quigley discloses “a cable modem system wherein information is

communicated between a plurality of cable modems and a cable modem termination system.” Quigley at 1:32-35; see also Abstract, 8:42-48 (Ex. 1014).

Figure 1 illustrates “a hybrid fiber coaxial (HFC) network showing typical pathways for data transmission between the headend (which contains the cable modem termination system) and a plurality of homes (each of which contains a cable modem).” Quigley at 3:57-61; see also 8:55-62, 10:41-48.



Quigley at Fig. 1. The '196 Patent similarly describes “a typical network as found in a cable television and/or video delivery network employing Hybrid Fiber-Coaxial (HFC) wiring scheme as disclosed in the prior art.” '196 Patent at 2:3-7, Fig. 1 (Ex. 1001).

125. In Quigley, “[t]he hybrid fiber coaxial network of a cable modem system utilizes a point-to-multipoint topology to facilitate communication between

the cable modem termination system and the plurality of cable modems.” Quigley at 9:1-4 (Ex. 1014). “Frequency domain multiple access (FDMA)/time division multiplexing (TDM) is used to facilitate communication from the cable modem termination system to each of the cable modems, i.e., in the downstream direction.” Quigley at 9:4-8. “Frequency domain multiple access (FDMA)/time domain multiple access (TDMA) is used to facilitate communication from each cable modem to the cable modem termination system, i.e., in the upstream direction.” Quigley at 9:8-12; see also 9:48-52 (“Contemporary cable modem systems operate on a plurality of upstream channels and utilize time division multiple access (TDMA) in order to facilitate communication between a plurality of cable modems and a single cable modem termination system on each upstream channel.”), 11:49-12:2, 37:29-45.

126. In describing the cable modem termination system upstream interface, Quigley explains that “cable modem transmit opportunities” each “define a time slot within which a designated cable modem 12 is permitted to transmit the data packet for which it previously sent a request to the cable modem termination system 10.” Quigley at 46:14-21; see also 9:56-10:3 (Ex. 1014). “The upstream channel 491, is divided into a plurality of time intervals 110, each of which may optionally be further subdivided into a plurality of sub-intervals.” Quigley at 46:31-34. “The upstream channel 491 is thus partitioned so as to facilitate the

definition of time slots, such that each of a plurality of cable modems 12 may transmit data packets to the cable modem termination system 10 without interfering with one another” Quigley at 46:34-40. Thus, in Quigley, the cable modem termination system located at the headend partitions the upstream channel into time slots associated with each of the cable modems. Quigley at 45:62-46:55, Fig. 36.

E. Overview of Banker

127. U.S. Patent No. 5,477,262, which issued to Banker and others, is entitled “Method and Apparatus for Providing an On-Screen User interface for a Subscription Television Terminal.” Banker at cover (Ex. 1015). It issued from U.S. Patent Application 07/800,836, which was filed on November 29, 1991. I understand that Banker is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

128. Banker generally describes “an improved user friendly interface for a subscription television terminal. In this manner, the first-time user should not be afraid to use their assigned subscription television terminal.” Banker at 3:19-22 (Ex. 1015). A “subscription television terminal” is simply another name for a “set top terminal.” Banker at 2:29. Banker then describes a number of user interface features for use on a subscription television terminal, such as providing “on-screen instructions to a user to assist them” in navigating the terminal, “visual

correspondence to keys of a user input device in such instructions, for example, by the use of brackets,” and “provid[ing] an impulse pay-per-view feature” in the terminal. See Banker at 4:1-5, 4:16-18. Banker also discussed how customers can be billed for using the subscription television terminal. See Banker at 7:58-8:3, 12:1-15.

F. Overview of Gordon

129. U.S. Patent No. 6,314,573, which issued to Gordon and others, is entitled “Method and Apparatus for Providing Subscription-On-Demand Services for an Interactive Information Distribution System.” Gordon at cover (Ex. 1016). It issued from U.S. Patent Application 09/086,799, which was filed on May 29, 1998. I understand that Gordon is prior art to the ’196 Patent. It was not cited during prosecution of the ’196 Patent.

130. Gordon generally describes a “method and apparatus for providing subscription-on-demand (SOD) services for a interactive information distribution system, where a consumer may subscribe to packages of on-demand programs for a single price and view the programs in the subscribed package at any time for no additional cost.” Gordon at Abstract (Ex. 1016). Gordon also describes “software” and “hardware” that can be used to implement its system and method, such as graphical user interfaces used in subscriber terminals for interacting subscription on demand services. Gordon at Abstract.

VIII. THE CHALLENGED CLAIMS ARE RENDERED OBVIOUS BY MURDOCK ALONE OR IN COMBINATION WITH NAZARATHY, QUIGLEY, BANKER, OR GORDON.

131. It is my opinion that claims 1-2, 4-6, 12-13, 27-28, 30-32, and 38-42 (the “challenged claims”) of the ’196 Patent are rendered obvious to a person of ordinary skill in the art based on Murdock alone. Murdock discloses a system for providing programming content in response to voice commands issued by users in their home. Murdock explains that voice commands from multiple users are multiplexed together and transmitted over a back channel to a remote server, which processes these commands to identify the requesting user and requesting content, and transmits the requested content back to the requesting user. Although Murdock does not explicitly state that the remote server demultiplexes the multiplexed signal prior to processing it, persons of ordinary skill in the art would have understood that demultiplexing would be required. Thus, as explained below, Murdock alone renders the challenged claims obvious.

132. In addition, Nazarathy and Quigley also disclose combining multiple signals from multiple locales into a single signal for transmission to a server, and then partitioning that combined signal into its constituent parts. The combination of Murdock with Nazarathy and Quigley thus also renders the challenged claims obvious.

133. Claims 5, 6, 31, and 32 of the ’196 Patent additionally require

generating “financial consequences” based on the content requested by the user. Murdock discloses implementing its system on pay-per-view, and Banker and Gordon provide further details regarding how pay-per-view systems work. Thus, Murdock alone and in combination with Banker and Gordon also render claims 5, 6, 31, and 32 obvious.

134. In the following paragraphs, I cite exemplary portions of the prior art references to show that each element of the challenged claims is disclosed by (i) Murdock alone; (ii) Murdock combined with Nazarathy; (iii) Murdock combined with Quigley; (iv) Murdock combined with Banker; and (v) Murdock combined with Gordon. In the section that follows, I discuss further the reasons why a person of ordinary skill in the art would have been motivated to combine Murdock with the teachings of Nazarathy, Quigley, Banker, and Gordon.

A. Every Element of the Challenged Claims Is Disclosed by Murdock Alone or Combined with Nazarathy, Quigley, Banker, or Gordon

1. Claim 1

135. The preamble of claim 1 reads as follows:

A method of using a back channel containing a multiplicity of identified speech channels from a multiplicity of user sites presented to a speech processing system at a wireline node in a network supporting at least one of cable television delivery and video delivery, comprising the steps of:

136. Murdock discloses the method recited in claim 1. Murdock uses a

“back channel containing a multiplicity of identified speech channels from a multiplicity of user sites” to deliver voice commands to a remote speech processing system at a wireline node, as recited in claim 1. In Murdock, users request programming using a verbal command. Murdock at 2:24-26, 2:47-49 (Ex. 1010). These commands are transmitted to a local processing unit 120 (such as a set top terminal or cable box). Murdock at 2:35-49. Then,

[t]he back channel multiplexer 136 multiplexes or combines the audio signal transmitted from the local processing unit 120 with additional audio signals from the local processing units of other users, i.e., other users of other voice activated home entertainment systems. The multiplexed audio signal is then transmitted via the back channel 134 to the service provider. The back channel 134 combines the multiplexed audio signal with additional multiplexed audio signals from other back channel multiplexers. As such, the back channel 134 transmits audio control signals from a plurality of local processing units 120 to the remote server computer 130 at the service provider.

Murdock at 3:1-12. After receiving and processing the multiplexed audio signals, the remote server 130 retrieves the requested programming content and transmits it to the requesting user’s local processing unit via forward channel 132. Murdock at 3:28-41. Murdock further discloses that the “network” supports delivery of “at least one of cable television delivery and video delivery,” as recited in the preamble of claim 1. Murdock at 2:28-31 (“Examples of the requested

programming content include web-based content, video-on-demand, cable television programming, and the like.”).

137. In Murdock, remote server computer 130 is a “wireline node” as that term is used in the ’196 Patent because it is a “a network node providing video or cable television delivery to multiple users over physical wires between the node and the multiple user locations.” See Paragraphs 80-81, above. For example, the remote server computer 130 provides video or cable television delivery to multiple users. See Murdock at 3:31-36 (“Once the user is identified and the requested programming content is determined, the server computer 130 retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132 to the local processing unit 120.”); see also 5:23-34 (Ex. 1010). The network the remote server 130 is connected to can provide cable television programming and interact with cable set top boxes. Murdock at 2:28-31 (“Examples of the requested programming content include web-based content, video-on-demand, cable television programming, and the like.”), 4:14-15 (“Examples of the local processing unit 120 may include set top boxes, cable boxes, and the like.”). Persons of ordinary skill in the art understand that cable television delivery and cable set top boxes operate on networks that are connected over physical wires, as the name “cable” suggests. Murdock’s back channel 134 is a “back channel” as that term is used in the patent because it is an

“upstream communication channel delivering signals from multiple user sites to a central wireline node” (i.e., remote server computer 130). See Paragraphs 82-83, above; Murdock at Abstract (“local processing unit ... transmits the received first audio signal to a service provider via the back channel” and the “remote server computer receives the first audio signal from the back channel.”); see also 3:15-17 (“[T]he remote server computer 130 receives the multiplexed signal from the back channel 134 and performs speech recognition on the received signal.”), 5:2-4 (“The command signal interface 402 [of the remote server computer] receives the multiplexed signal from the back channel 134.”). Thus, Murdock discloses the preamble of claim 1 of the ’196 Patent: Murdock discloses a back channel containing a multiplicity of identified speech channels (the audio signals received from multiple local processing units of users) from a multiplicity of user sites (the locations of the local processing units of the users), which presents such speech channels via a wireline node to a speech processing system (the remote server computer at the service provider).

138. The same subject matter is also disclosed by the Murdock Provisional. Like Murdock, the Murdock Provisional also discloses a “backward channel multiplexor” that combines audio signals (consisting of speech commands issued from several users) from several “home entertainment processors” and sends the combined audio signal to the “Information Source.” Murdock Provisional at 1:15-

19, 2:16-22, 3:28-32 (Ex. 1011). The Information Source is “located at the site of the information provider,” which can provide “web content, video-on-demand, or cable television.” Murdock Provisional at 4:1-4. It also includes a speech recognition server. Murdock Provisional at 4:3-6. Like the remote server computer of Murdock, the central speech recognition server at the Information Source processes the audio signals received from the backward channel multiplexor and sends the appropriate programming modifications to the users via a forward channel. Murdock Provisional at 1:18-22.

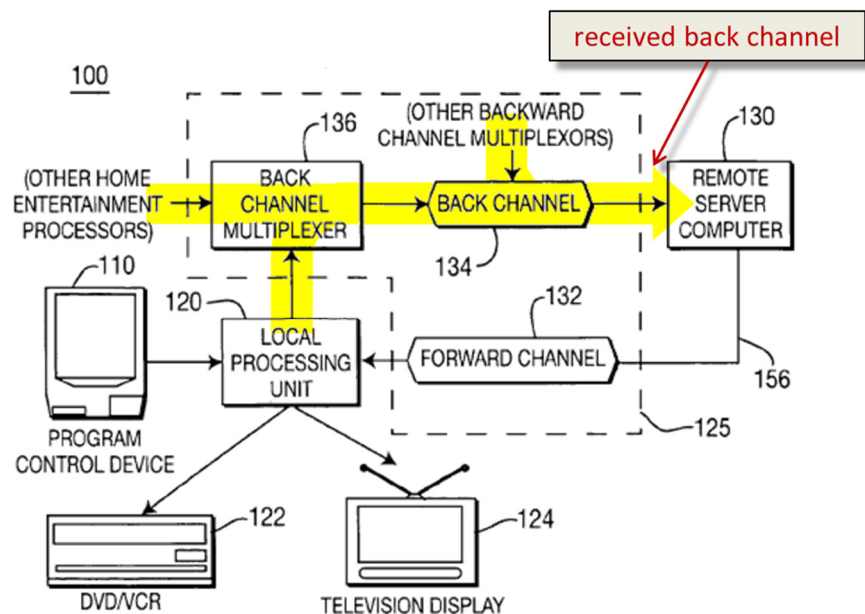
139. The first element of claim 1 reads as follows:

receiving said back channel to create a received back channel;

140. Murdock discloses this element. In particular, Murdock states that the “local processing unit ... transmits the received first audio signal to a service provider via the back channel” and the “remote server computer receives the first audio signal from the back channel.” Murdock at Abstract (Ex. 1010); see also 3:15-17 (“[T]he remote server computer 130 receives the multiplexed signal from the back channel 134 and performs speech recognition on the received signal.”), 5:2-4 (“The command signal interface 402 [of the remote server computer] receives the multiplexed signal from the back channel 134.”). Murdock further explains that the “back channel multiplexer 136 multiplexes or combines” the signal from local processing unit 120 “with additional audio signals from the local

processing unit of other users, i.e., other users of other voice activated home entertainment systems.” Murdock at 3:1-5. The back channel also “combines the multiplexed signals from other back channel multiplexers.” Murdock at 3:7-9.

141. Figure 1 of Murdock illustrates the combined signals transmitted from multiple different users over back channel 134 to remote server computer 130, which receives the back channel.



Murdock at Fig. 1 (Ex. 1010) (annotated). Thus, Murdock discloses remote server computer 130 receiving a “back channel to create a received back channel,” as recited in claim 1.

142. The Murdock Provisional also discloses this subject matter. For example, the Murdock Provisional states:

It is the second advantage of this invention to provide a mechanism, the network’s backward channel, for transmitting voice control data to

the network information source. The central speech recognition server is contained at the network information source. It processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.

Murdock Provisional at 1:16-22 (Ex. 1011). The Murdock Provisional further provides that “[t]he back channel is the channel in which the Voice Control Data from the HEP is transmitted to the information source.” Murdock Provisional at 2:9-11. In other words, the information source receives the back channel containing voice control data to create a received back channel. This is illustrated in Figure 1 of the Murdock Provisional is essentially the same as Figure 1 of Murdock. It shows the Information Provider, which contains the Information Source (Murdock Provisional at 4:1-2), receiving the back channel to create a received back channel. Murdock Provisional at Fig. 1.

143. The next element of claim 1 reads as follows:

partitioning said received back channel into a multiplicity of received identified speech channels;

144. Murdock discloses this element. As discussed above, Murdock discloses transmitting a multiplexed signal (containing the combined voice signals of multiple different users) over back channel 134 to remote server computer 130. Murdock at 3:1-12 (Ex. 1010). Murdock discloses that once these voice

commands are received via the back channel, the remote computer then “identif[ies] and recogniz[es] the user that generated the audio signal and determin[es] the requested programming content contained in the audio signal.” Murdock at 3:18-31. This would require demultiplexing the multiplexed signal, as would be apparent to a person of ordinary skill in the art though Murdock does not explicitly refer to “demultiplexing” the received combined signal.

145. A person of ordinary skill in the art would have known that to make any sense of the multiplexed signal, such as using it to identify or recognize a specific user and to determine the programming the particular user requested, the audio signal must first be demultiplexed (which constitutes “partitioning,” as explained in Paragraphs 84-89, above). This goes hand-in-hand with multiplexing—if a signal is multiplexed (i.e., combined) on one end for transmission, then the signal must be demultiplexed (i.e., partitioned) on the other end, so that that multiplexed signal can be interpreted and used. See, e.g., Nazarathy at 14:62-64 (Ex. 1013); Quigley at 51:31-35 (“The signals of frame 179 from different cable modems 12a, 12b, 12c, 12d, etc. ... are introduced in upstream data processing through upstream channel 191 ... to a TDMA demultiplexer 192 ... in the cable modem termination system”) (Ex. 1014). Receiving a multiplexed signal without de-multiplexing it would be akin to trying to simultaneously listen to many people talking at the same time. Thus, the remote server computer in

Murdock must demultiplex the multiplexed signal prior to identifying or recognizing the user that generated a particular audio signal. These demultiplexed signals are the speech commands sent by the individual users via their respective local processing units and constitute the multiplicity of “identified speech channels” recited in this element.

146. Thus, Murdock discloses to a person of ordinary skill in the art that remote server 130 demultiplexes the received back channel into multiple portions that are each identified with speech commands originating from the various different system users. This constitutes “partitioning said received back channel into a multiplicity of received identified speech channels,” as recited in claim 1. See Paragraphs 84-89, above. Murdock alone discloses the “partitioning” step of claim 1.

147. The Murdock Provisional also discloses the subject matter of Murdock described with respect to this element. Specifically, the provisional application discloses a “Backward Channel Multiplexer” that “mixes the backward channel content received from several HEPs into one signal that is sent to the Information source” as is illustrated in Figure 1. Murdock Provisional at 1:18-19, 3:18-30, Fig. 1 (Ex. 1011). The central speech recognition server at the Information Source “processes speech control data arriving through the backward channel”—i.e., the multiplexed audio signal—“and then sends the appropriate

programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22. For the reasons discussed above, a person of ordinary skill in the art would have understood that in order to perform speech recognition on the viewer’s speech control data, the speech recognition server must first demultiplex the multiplexed signal into its component parts—the individual signals sent by each of the individual users). Thus, Murdock alone discloses the “partitioning” step of claim 1, and the Murdock Provisional discloses the same subject matter.

148. Further, Murdock can be combined with Nazarathy to disclose “partitioning” a received back channel as recited in claim 1 of the ’196 Patent. In particular, Nazarathy discloses a “hybrid coaxial cable network” in which data signals in the “return path” (i.e., the back channel) are multiplexed for transmission to the headend. Nazarathy at Abstract, 13:64-14:46 (Ex. 1013). For instance, Nazarathy discloses “TDM multiplexing” (i.e., time division multiplexing) and “WDM multiplexing” (i.e., wavelength division multiplexing) the signals on the return path. Nazarathy at 14:6-8 (“There are several possibilities, e.g., WDM multiplexing, TDM multiplexing, or a combination of both WDM and TDM multiplexing”); see also 14:9-61. Nazarathy then expressly discloses that the multiplexed signal is demultiplexed at the cable headend: “Any operations of TDM and/or WDM multiplexing are undone at the HE [cable headend] by

corresponding WDM and TDM demultiplexers.” Nazarathy at 14:62-64. Thus, Nazarathy teaches that the multiplexed back channel of Murdock would be demultiplexed at the remote server computer to identify the incoming data signals.

149. As discussed in more detail below, a person of ordinary skill in the art would have been motivated to combine Murdock and Nazarathy in this manner. See Paragraphs 294-299, below.

150. Further, Murdock can be combined with Quigley to disclose “partitioning” a received back channel as recited in claim 1 of the ’196 Patent. In particular, Quigley describes a cable network in which the “upstream channel” is “partitioned” by dividing the channel into a plurality of time intervals. Quigley at 46:32-40 (Ex. 1014). This partitioning into “time slots” is necessary to “accomplish TDMA for upstream communication,” as disclosed in Quigley. Quigley at 9:56-10:3 (“In order to accomplish TDMA for upstream communication, it is necessary to assign time slots within which cable modems having a message to send to the cable modem termination system are allowed to transmit.”). As disclosed in Quigley, the cable modem termination system located at the headend partitions the upstream channel into time slots associated with each of the cable modems. Quigley at 9:56-10:3, Thus, the combination of Murdock with the TDMA upstream communication technique of Quigley discloses “partitioning said received back channel into a multiplicity of received identified

speech channels,” as recited in claim 1.

151. As discussed in more detail below, a person of ordinary skill in the art would have been motivated to employ the TDMA upstream communication technique of Quigley in the “voice activated system” of Murdock. See Paragraphs 294-299, below.

152. The next element of claim 1 reads as follows:

processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content; and

153. Murdock discloses this element. After receiving and partitioning the multiplexed signal from the back channel 134, the remote server computer 130 “performs speech recognition” on the received signals. Murdock at 3:13-17 (Ex. 1010). This is done by “identifying or recognizing the user that generated the audio signal and determining the requested programming content contained in the audio signal.” Murdock at 3:28-31. Murdock further confirms that after receiving the signals, “the CPU 406 [of the remote server computer] executes the speech recognition module 410,” and the “the remote server computer 130 recognizes the user requesting the programming content and the request from the received signal.” Murdock at 5:11-14. Thus, the remote server computer 130 performs “processing” of the “multiplicity of received identified speech channels” (i.e., voice commands from multiple different users) to “create a multiplicity of identified speech content”

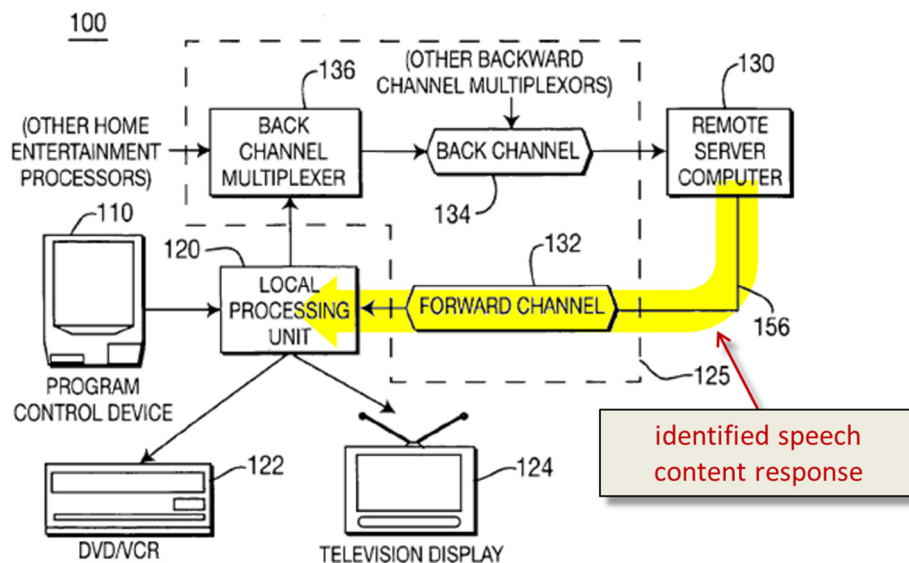
(i.e., the identity of the requested content), as recited in claim 1.

154. The Murdock Provisional also discloses this subject matter. Specifically, the Murdock Provisional states that the central speech recognition server, located at the Information Source, “processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011). To “process[]” the speech control data to “send[] the appropriate programming modifications through the forward channel as requested by the viewer,” the central speech recognition server must identify the requested programming.

155. The final element of claim 1 reads as follows:

responding to said identified speech content to create an identified speech content response that is unique, for each of said multiplicity of identified speech contents.

156. Murdock discloses this element of claim 1. In Murdock, after speech processing is performed to recognize the spoken request for particular content (i.e., said identified speech content), “the server computer 130 retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132” to the requesting user. Murdock at 3:31-36, 5:23-34 (Ex. 1010).



Murdock at Fig. 1 (annotated). In other words, the server computer “responds” to the identified speech content by retrieving the requested programming content and transmitting it to the particular viewer who requested the particular content. Thus, Murdock responds to each voice command by creating “an identified speech content response” (requested programming content) that is “unique” in that it responds to the specific spoken request of a particular viewer. Murdock thus discloses this element of claim 1.

157. The Murdock Provisional also discloses the same subject matter. As discussed above, the Murdock Provisional discloses identifying the content requested by the user. After the requested programming for each viewer is identified, the central speech recognition server “then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011).

2. Claim 2

158. Claim 2 is a dependent claim that adds several limitations. Claim 2 begins as follows:

The method of claim 1, wherein at least one of said identified speech channels has an associated user site; and

159. Murdock discloses this limitation. As discussed in connection with claim 1, the “identified speech channels” are the audio signals (i.e., voice commands) issued from individual local processing units and received by the remote server computer via back channel 134. Murdock at Abstract, 2:44-67, 3:15-17, 5:2-4 (Ex. 1010). The local processing unit can be a “set top terminal, a cable box, and the like.” Murdock at 2:44-47. This type of equipment is located at a particular user site. Indeed, this is how the ’196 Patent describes user sites: “Each user site contains a Set Top Box, such as STB 180,” ’196 Patent at 2:8; see also 4:18-19 (“As in FIG. 1, each user site contains a Set Top Box, such as STB 180”) (Ex. 1001). Because each identified speech channel comes from a local processing unit, which is located at a user site, “at least one of said identified speech channels has an associated user site,” as recited in claim 2.

160. The Murdock Provisional also discloses this subject matter. The Murdock Provisional discloses that individual users send speech signals from a hand-held controller and home entertainment processor to the central speech

recognition server via a back channel. Murdock Provisional at 2:1-3:2 (Ex. 1011).

161. Claim 2 then adds the following additional limitation:

**wherein the step partitioning said received back channel is
further comprised of the step of:
partitioning said received back channel into a multiplicity of said
received identified speech channels from said associated user site;**

162. Murdock, alone or in combination with Nazarathy and Quigley, discloses this limitation. This part of claim 2 is similar to the “partitioning” step of claim 1 but includes the additional limitation that the received identified speech channels be “from said associated user site.” In Murdock, audio signals (i.e., voice commands) are transmitted from individual local processing units to the remote server computer via back channel 134 (i.e., the “received back channel”). Murdock at Abstract, 2:44-67, 3:15-17, 5:2-4 (Ex. 1010). As discussed in connection with the “partitioning” step of claim 1, the multiplexed signal received via back channel 134 would be demultiplexed (i.e., partitioned) to identify signals from individual users, as would be apparent to a person of ordinary skill in the art. See Paragraphs 144-146, above. This is implicit in the disclosure of Murdock, and explicitly disclosed by Nazarathy and Quigley. Thus, the “received back channel” (i.e., back channel 134) is partitioned “into a multiplicity of said received identified speech channels from said associated user site” (i.e., local processing unit 120), as recited in the claim.

163. Each “received identified speech channel” from a particular “associated user site” may be further partitioned as well. For instance, a viewer may speak multiple, separate voice commands. As suggested by Murdock, the user can speak the command “Show me all the college football games on this Saturday.” Murdock at 3:53-55 (Ex. 1010). The user could then review the list of options and speak a command selecting a particular game to view (or record). In this case, the signal from the user site would be partitioned based on the two separate spoken commands.

164. Further, a person of ordinary skill would also recognize that even the individual signals from each local processing unit are commonly divided into discrete portions for transmission. Quigley expressly discloses that, in one embodiment, the cable modem “fragments the data to be transmitted into a plurality of segments if the requested amount is larger than the granted [bandwidth] amount.” Quigley at 1:65-2:1 (Ex. 1014). In Quigley, the segments are then transmitted to the headend using the disclosed TDMA technique. Quigley at 2:2-29, 9:56-10:3; see also 46:32-40 (“The upstream channel 491, is divided into a plurality of time intervals 110, each of which may optionally be further subdivided into a plurality of sub-intervals 489. The upstream channel 491 is thus partitioned so as to facilitate the definition of time slots, such that each of a plurality of cable modems 12 may transmit data packets to the cable modem

termination system 10 without interfering with one another, e.g., without having data collisions due to data packets being transmitted at the same time.”). The headend then reassembles the segments to create each particular signal “from said associated user site.” In order to reassemble the segments to create the particular signal from said associated user site, the headend must first divide the received segments into multiple portions that are each identified specific user sites, and then reassemble those portions that are identified with the same user site together. In this process, the headend “divides” the “received back channel” (the received segments) into “multiple portions that are each identified” with data “originating from system users,” and thus satisfies the construction of “partitioning” as that term is used in the ’196 Patent. See Paragraph 87.

165. Thus, Murdock alone or in combination with either Nazarathy or Quigley discloses “partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site,” as recited in claim 2.

166. The Murdock Provisional also discloses this subject matter. The Murdock Provisional discloses a “Backward Channel Multiplexer” that “mixes the backward channel content received from several HEPs into one signal that is sent to the Information source.” Murdock Provisional at 3:28-30 (Ex. 1011). Thus, the Backward Channel Multiplexer multiplexes several audio signals into one and

sends it to the Information Source, which, as discussed above, contains the central speech recognition server. Murdock Provisional at 1:18-19; 3:18-30.

167. The central speech recognition server “processes speech control data arriving through the backward channel”—i.e., the multiplexed audio signal—“and then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011). To be able to “send[] the appropriate programming modifications through the forward channel” as requested by each viewer, the speech recognition server at the information source would first demultiplex the multiplexed signal containing the “speech control data” from each viewer. Because the speech control data from various users each come from an individual HEP, they are each “associated” with a user site—namely, the location of the HEP the audio signal originated from. The Murdock Provisional also discloses that audio signals from a single HEP location can be multiplexed and demultiplexed multiple times. This is because every time the viewer at the user site sends an audio signal, the audio signal is multiplexed and demultiplexed, as described above.

168. Claim 2 then adds the following additional limitation:

wherein the step processing said multiplicity of said received identified speech channels is further comprised the step of: processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of

said identified speech contents; and

169. Murdock discloses this limitation. After receiving and partitioning the multiplexed signals via back channel 134, the remote server computer 130 “performs speech recognition” on the demultiplexed signals. Murdock at 3:13-17 (Ex. 1010). This is done by “identifying or recognizing the user that generated the audio signal and determining the requested programming content contained in the audio signal.” Murdock at 3:28-31. Murdock further confirms that after receiving the demultiplexed signals, “the CPU 406 [of the remote server computer] executes the speech recognition module 410, [and] the remote server computer 130 recognizes the user requesting the programming content and the request from the received signal.” Murdock at 5:11-14.

170. Thus, the remote server computer “processes” the “multiplicity of said received identified speech channels from said associated user site” (i.e., audible signals transmitted by local processing units) by performing speech recognition on each of these channels “to create a multiplicity of said identified speech contents” (i.e., the identity of the requested programming). These speech channels are also from “associated user site[s]”, as discussed in connection with the first element of claim 2. Paragraph 159, above; Murdock at 2:44-47.

171. The Murdock Provisional also discloses this subject matter. The Murdock Provisional states that the central speech recognition server, located at

the Information Source, “processes speech control data arriving through the backward channel”—i.e., the multiplexed audio signal—“and then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011). In other words, the central speech recognition server first identifies the requested programming, and then sends that requested programming to the viewer at the user site. Moreover, the central speech recognition server “processes” a multiplicity of speech control data content from a single user because it processes the speech control data every time the user uses it to request content.

172. Claim 2 then adds the following additional limitation:

**wherein the step responding to said identified speech content is further comprised the step of:
responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.**

173. Murdock discloses this limitation. After performing speech recognition on the user’s voice command (i.e., the identified speech content), “the server computer 130 retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132 to the local processing unit 120” that requested the programming content.

Murdock at 3:31-36, 5:23-34 (Ex. 1010). Thus, the remote server computer

“responds” to the identified speech content for a specific user site (the location of the local processing unit 120) to “create said identified speech content response” (the requested content) and transmits it back to the “associated user site,” as recited in the claim.

174. The Murdock Provisional also discloses this subject matter. As discussed above, the Murdock Provisional describes processing the speech control data to identify the user-requested programming content. After the requested programming for each viewer is identified, the central speech recognition server “then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011).

3. Claim 4

175. Claim 4 is a dependent claim that adds the following limitation:

**The method step of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
processing said identified speech content response to create said identified user site response; and sending said identified user site response to said identified user site.**

176. Murdock discloses this added limitation. In particular, Murdock discloses the that “remote computer server 130 performs the speech recognition by identifying or recognizing the user that generated the audio signal and determining

the requested programming content contained in the audio signal.” Murdock at 3:28-31 (Ex. 1010). Then “the server computer 130 retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132 to the local processing unit 120” of the viewer that requested the programming. Murdock at 3:31-36.; see also 5:23-34. Thus, Murdock discloses processing the “identified speech content response to create said identified user site response” by retrieving the requested program content from a program database and preparing it for transmission and then “sending said identified user site response to said identified user site” (i.e., delivering the requested content to local processing unit 120 for display), as recited in the claim.

177. The Murdock Provisional also discloses this subject matter. After the requested programming for each viewer is identified by the Information Provider, the central speech recognition server “then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22 (Ex. 1011). For instance, the Information Provider can identify and retrieve “web content, video-on-demand, or cable television” programming to send to the requesting Home Entertainment Processor (HEP). Murdock Provisional at 1:18-22, 2:5-14.

4. Claim 5

178. Claim 5 is a dependent claim that adds the following limitation:

**The method of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and**

179. Murdock discloses this limitation. In particular, Murdock discloses requesting content, which can include video-on-demand or pay-per-view, and then processing that request. See Murdock at 2:24-31 (“The input command signal may comprise a verbal request of programming content from a service provider....Examples of the requested programming content include web-based content, video-on-demand, cable television programming, and the like.”); see also 1:16-37 (discussing the need in the art “for a system and a concomitant method for economically provided video content in response to a spoken command” to address pay-per-view scenarios, among others), 3:31-36 (Ex. 1010). A person of ordinary skill in the art would have known that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. Thus, Murdock discloses “assessing said speech response” (i.e., assessing the content requested by the user) associated with a particular “user device” (i.e., program control device or local processing unit) to “create a financial consequence (i.e., payment for a pay-per-view or video-on-demand movie), as recited in the claim.

180. The Murdock Provisional also discloses this same subject matter I have cited in the Murdock patent for this claim. Murdock Provisional at 2:12-14 (noting that the requested content can include “video-on-demand”), 4:1-3 (same); 2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex. 1011).

181. The last step of claim 5 requires:

billing said user site based upon said financial consequence.

182. Murdock discloses this limitation. As discussed above, Murdock discloses requesting content, which can include video-on-demand or pay-per-view. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). Also as mentioned above, a person of ordinary skill in the art would have known that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill.

183. Murdock alone discloses the limitations of claim 5. In addition, Murdock could be combined with the teaching of numerous other prior art references that disclose pay-per-view movie functionality in cable and other television networks. By the end of the 1990s, pay-per-view systems had been implemented by a number of different cable operators and the concept was well known. A person of ordinary skill in the art would have been familiar with the basic functionality of a pay-per-view system, which generally follows the steps

recited in claim 5 (and also claims 6, 31, and 32) of the '196 Patent. Specifically, a typical prior art pay-per-view movie system would present the user with a list of available titles and either the associated pricing or at least a warning that selecting a title from the list would lead to a charge—hence the name “pay-per-view.”

When the user selected a movie to watch, the system would commonly ask the user to confirm the purchase (for example by pressing a particular button on the television remote control). Once the user confirmed the purchase, the pay-per-view system would deliver the requested movie for viewing and bill the user, commonly by adding the purchase price to the subscriber's monthly bill.

184. A person of ordinary skill in the art would have understood these basic features of a pay-per-view system, like the one disclosed in Murdock, because they were well known in the prior art. For example, U.S. Patent No. 5,477,262 is a prior art patent entitled “Method and Apparatus for Providing an On-Screen User Interface for a Subscription Television Terminal” that issued in 1995. Banker at cover (Ex. 1015). The Banker prior art patent describes the basic steps of permitting a user to purchase pay-per-view content. Banker at 16:62-17:3, Figs. 6F. In the Banker system, a user can purchase a pay-per-view movie by pressing the “BUY key” on the television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Figs. 6F; see also 16:51-55 (“When the terminal of FIG. 3 is in the On mode

and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by actuating the BUY key initiates a buy sequence. Consequently, this key is used to purchase an event.”). Once entered by the user, the “code is checked with the code stored in memory” (i.e., “assessing a response identified” with the user, as recited in claim 8). “If correctly entered, the subscriber is asked to press the BUY key again ... and the subscriber is thanked” Banker at 17:1-3. The user is then billed for the purchase. Banker at 7:60-63 (“Billing computer 101 includes a subscriber database and generates a monthly bill for the subscribers in the system based on level of service and any pay-per-view and impulse pay-per-view purchases.”). Thus, Banker discloses the “assessing” and “billing” steps of claim 5. Moreover, as explained below in paragraphs 300-301, a person of ordinary skill in the art would have been motivated to combine Banker with Murdock. Accordingly, Murdock alone or in combination with Banker discloses the elements of claim 5 and renders the claim obvious.

185. As another example, U.S. Patent No. 6,314,573 is a prior art patent entitled “Method And Apparatus For Providing Subscription-On-Demand Services For An Interactive Information Distribution System” that also describes the basic steps of permitting a user to purchase pay-per-view content using a different platform. Gordon at 8:66-9:63, Fig. 3B (Ex. 1016). The Gordon patent was

assigned to Diva Systems, which offered a subscription pay-per-view platform in the 1990s. The Gordon system offers the subscriber an option to purchase an on-demand program subscription. Gordon at 9:64-10:9, Figs. 3B and 8. If the user elects to make a purchase, the system requests a master PIN as a confirmation. Gordon at 10:16-27. If the PIN number is valid, the system updates the billing system with the new subscriber's account number. Gordon at 10:43-46. ("After updating the subscription data base, the routine updates, at step 382, the administrative system such that the interactive information distribution system's billing system is updated with the new subscriber's account number."); Gordon at 2:60-63 ("As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that package and is billed accordingly."). Once a user subscribes to a package of programming and pays the fixed price, he can view any program within the package at any time. Gordon at 3:61-63. Moreover, as explained below in paragraphs 300-301, a person of ordinary skill in the art would have been motivated to combine Gordon with Murdock. Accordingly, Murdock alone or in combination with Gordon discloses the elements of claim 5 and renders the claim obvious.

186. The Murdock Provisional also discloses this same subject matter I have cited in the Murdock patent for this claim. Murdock Provisional at 2:12-14 (noting that the requested content can include "video-on-demand"), 4:1-3 (same);

2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex. 1011).

5. Claim 6

187. Claim 5 is a dependent claim that adds the following limitation:

**The method of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
assessing said speech response to create a financial consequence identified with said user site;**

188. As discussed in connection with claim 5, Murdock discloses that a viewer can use voice commands to request content, which can include video-on-demand or pay-per-view. Murdock at 2:24-31; see also 1:16-37 (Ex. 1010). Murdock also discloses processing that video-on-demand or pay-per view request. See Murdock at 3:31-36. A person of ordinary skill in the art would have known that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. Thus, Murdock discloses “assessing” a response (i.e., assessing the content requested by the user) associated with a particular “user device” (i.e., program control device or local processing unit) to “create a financial consequence (i.e., payment for a pay-per-view or video-on-demand movie), as recited in the claim.

189. The Murdock Provisional also discloses this same subject matter.

Murdock Provisional at 2:12-14 (noting that the requested content can include “video-on-demand”), 4:1-3 (same); 2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex. 1011).

190. In addition, Banker also discloses this limitation of the claim. In Banker, a user can purchase a pay-per-view movie by pressing the “BUY key” on the television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (“When the terminal of FIG. 3 is in the On mode and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by actuating the BUY key initiates a buy sequence. Consequently, this key is used to purchase an event.”) (Ex. 1015). Thus, Banker discloses “assessing” a response (i.e., assessing the user’s key entries) to “create a financial consequence” (i.e., purchase of a pay-per-view movie), as recited in the claim. The combination of Murdock with Banker also discloses this element.

191. Gordon also discloses this limitation of the claim. Gordon’s Figure 3B illustrates a flow chart of the basic steps in ordering video content. The user selects a category from a category menu, and then a title from a title menu (e.g. Sesame Street). Gordon at 9:18-28; Fig. 3B (Ex. 1016). After receiving a title selection, the system determines the title price. Gordon at 9:29-31; Fig. 3B. Hence, the system assesses the user’s response (i.e. choice of title) to create a

financial consequence, i.e. the title price. Thus, Gordon discloses “assessing” a response (i.e., assessing the user’s choice of a title) to “create a financial consequence” (i.e., determine the price of the title), as recited in the claim. The combination of Murdock with Gordon also discloses this element.

192. The next step of claim 6 requires:

displaying said financial consequence to create a displayed financial consequence at said user site;

193. As discussed in connection with claim 5, Murdock explains that a viewer can use voice commands to select and order pay-per-view movies. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). The Murdock Provisional also discloses this same subject matter. Murdock Provisional at 2:12-14 (noting that the requested content can include “video-on-demand”), 4:1-3 (same); 2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex. 1011).

194. A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase, such as by displaying the price of the requested programming to the user. For example, the Banker patent describes a prior art pay-per-view system that presented the user with a screen stating: “NEXT, PLEASE PRESS [BUY] TO PURCHASE THIS EVENT” to confirm the viewer’s intent to accept the purchase (i.e., the “financial

consequence”). Banker at 16:62-17:3, Fig. 6F (Ex. 1015). As another example, the Gordon patent illustrates at Fig. 3B a flow chart of the basic steps in ordering video content. Once the user selects a title and the price has been determined, the system queries the user to see if he wants to pay “a la carte” to view his selection. 9:56-63 (“...the video session manager sends, at step 352, an applet for a second type of title information screen. At step 354, the terminal creates a screen informing the subscriber of the reason for the non-zero price and also presents the user with an “a la carte” purchase option. The subscriber will then be able to purchase and view the a la carte selection or return to a previous menu.”) Thus, either Murdock alone or combined with Banker or Gordon discloses this limitation of claim 6.

195. Next, claim 6 requires:

confirming said displayed financial consequence from said user site to create a financial commitment; and

196. As discussed in connection with claim 5, Murdock explains that a viewer can use voice commands to select and order pay-per-view movies. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). The Murdock Provisional also discloses this same subject matter. Murdock Provisional at 2:12-14 (noting that the requested content can include “video-on-demand”), 4:1-3 (same); 2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex.

1011).

197. A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase and confirming their intent to do so. For example, the Banker patent describes a prior art pay-per-view system that requires the user to press the BUY key, enter a system provided code, and then asking the subscriber to “press the BUY key again” to confirm the intent to purchase (i.e., the “financial consequence”) to “create a financial commitment,” as claimed. Banker at 16:62-17:3, Fig. 6F (Ex. 1015). As noted above, the Gordon patent allows a viewer to pay “a la carte” to view a selection. Gordon at 9:56-63 (“...the video session manager sends, at step 352, an applet for a second type of title information screen. At step 354, the terminal creates a screen informing the subscriber of the reason for the non-zero price and also presents the user with an “a la carte” purchase option. The subscriber will then be able to purchase and view the a la carte selection or return to a previous menu.”) (Ex. 1016). Gordon provides an example in Figure 8, which gives the viewer the option to purchase a program subscription by clicking on-screen buttons. Gordon at Fig. 8; see also 6:43-47 (on-screen buttons are selected via the remote control). Thus, either Murdock alone or combined with Banker or Gordon discloses this limitation of claim 6.

198. The following figure is from Banker:

SCREEN 7a
IPPV BUY

NEXT, PLEASE PRESS
[BUY]
TO PURCHASE THIS EVENT

OR

PRESS [MENU] TO EXIT.

Banker at Fig. 6F (Ex. 1015) (excerpted).

199. The following figure is from Gordon:

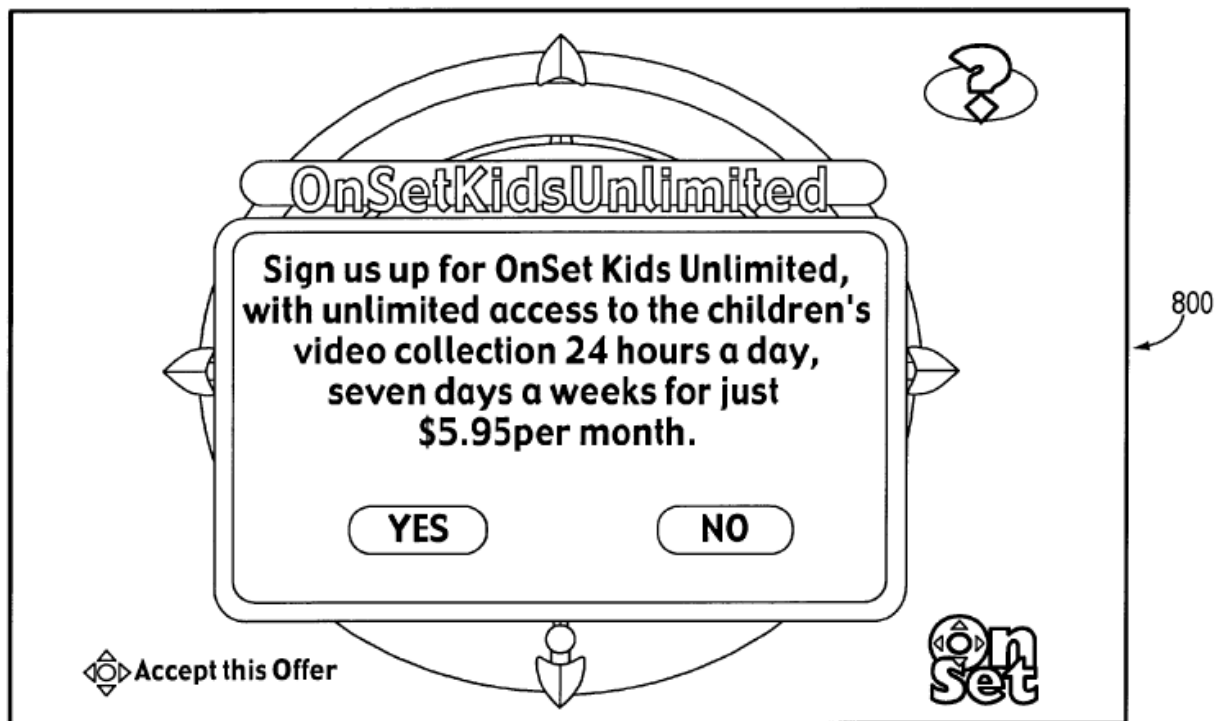


FIG. 8

Gordon at Fig. 8 (Ex. 1016).

200. Claim 6's last requirement is:

billing said user site based upon said financial commitment.

201. Murdock discloses this limitation. As discussed above, Murdock discloses requesting content, which can include video-on-demand or pay-per-view. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). The Murdock Provisional also discloses this same subject matter. Murdock Provisional at 2:12-14 (noting that the requested content can include "video-on-demand"), 4:1-3 (same); 2:19-33 (describing how speech commands for video-on-demand can be processed) (Ex. 1011).

202. Also as discussed above, a person of ordinary skill in the art would have known that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. Banker also discloses billing the subscriber for a pay-per-view purchase. Banker at 7:60-63, 16:62-17:3 ("generate a monthly bill ... based on level of service and any pay-per-view and impulse pay-per-view purchases") (Ex. 1015). Gordon also discloses billing the subscriber for a pay-per-view purchase. Gordon at 10:39-47 ("If the PIN is valid [] the routine updates, at step 382, the administrative system such that the interactive information distribution system's billing system is updated with the new subscriber's account number."); 2:60-63 ("As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that

package and is billed accordingly.”) (Ex. 1016). Thus, Murdock alone or combined with Banker or Gordon discloses “billing said user site based upon said financial commitment,” as claimed.

6. Claim 12

203. Claim 12 is a dependent claim that adds the following limitation:

The method of claim 2, wherein the step responding to said identified speech content from said associated user site is further comprised of the step of:

responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech content identified with said user site.

204. Murdock discloses this added limitation. The claim recites responding to the speech content identified as to said user site (i.e., the requested content) “based upon natural language.” ’196 Patent at 52:53 (Ex. 1001).

Murdock discloses performing speech recognition “based upon natural language,” as recited in the claim. In Murdock, users request programming content by issuing spoken commands. Murdock at 1:41-49, 2:22-26, 3:42-57 (Ex. 1010). Murdock provides the specific example of a user stating: “Show me all the college football games on this Saturday.” Murdock at 3:54-57. A person of ordinary skill in the art would recognize this as a “natural language” command because it reflects a natural manner of speaking.

205. In addition, Murdock explains that after identifying the programming content requested by a particular user (i.e., “speech content identified as to said user site”) based on a natural language request, server computer 130 “retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132 to the local processing unit 120” that requested the programming content. Murdock at 3:31-36; see also 5:23-34 (Ex. 1010). The server computer thus “responds” to the speech content identified for each user site (i.e., the location of the local processing unit 120) by creating a response (i.e., retrieving and transmitting the requested content) to the specific user site, as recited in the claim.

206. Thus, Murdock discloses “responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech content identified with said user site,” as recited in claim 12.

207. The Murdock Provisional also discloses this subject matter. It provides a similar natural language spoken command: “Show me all of the NCAA football games on this afternoon.” Murdock Provisional at 2:19-20 (Ex. 1011). It also explains that after the requested programming is identified, the central speech recognition server “then sends the appropriate programming modifications through the forward channel as requested by the viewer.” Murdock Provisional at 1:18-22.

The requested content is transmitted to the requesting user site.

7. Claim 13

208. Claim 13 is a dependent claim that adds the following limitation:

The method of claim 2, wherein the step processing said multiplicity of said received speech channels is further comprised for at least one of said user sites is further comprised of the step of:

processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.

209. Murdock discloses this added limitation. As discussed in connection with claim 12, Murdock discloses speech recognition processing “based upon natural language.” See Paragraph 204, above. In particular, Murdock provides an example of a natural language spoken command: “Show me all the college football games on this Saturday.” Murdock at 3:54-57 (Ex. 1010). A person of ordinary skill in the art would recognize this as a “natural language” command because it reflects a natural manner of speaking.

210. Further, Murdock discloses processing the spoken natural language command “to create said speech content identified as to said user site,” as recited in the claim. Murdock discloses that remote server computer 130 identifies the user and performs speech recognition to identify the requested content. Murdock at 3:13-17, 3:28-31, 5:11-14 (Ex. 1010). The identity of the requested content

constitutes the claimed “speech content identified as to said user site.”

211. The Murdock Provisional also discloses this subject matter. The Murdock Provisional discloses identifying the content requested by a user via speech commands. Murdock Provisional at 1:18-22 (Ex. 1011). It also provides an example of a natural language command the viewer can speak: “Show me all of the NCAA football games on this afternoon.” Murdock Provisional at 2:19-20.

8. Claim 27

212. The preamble of claim 27 reads as follows:

A system supporting speech recognition in a network, said system comprising:

213. Murdock discloses a system supporting speech recognition in a network as recited in claim 27. Murdock discloses “a system . . . for providing video content in response to an audio signal such as a spoken audio command” that includes a “remote server computer 130” which “performs . . . speech recognition.” Murdock at 1:41-43, 3:28-31 (Ex. 1010). The remote server computer 130 includes a “memory 404” that “stores a speech recognition module 410 that, when retrieved and executed by the CPU 406 [in the remote server 130], causes the remote server computer 130 to operate as a speech recognition server.” Murdock at 4:65-5:10. The remote server computer is part of a network that can provide “video-on-demand” and “cable television programming, and the like” to multiple

users. Murdock at 2:28-31; Fig. 1. Specifically, the network is a “television signal delivery network 125 comprising a forward channel 132, a back channel 134 and a back channel multiplexer 136.” Murdock at 2:19-21. The remote server computer, which includes the speech recognition module 410, is therefore a “ system supporting speech recognition in a network” as the preamble of claim 27 requires.

214. The Murdock Provisional also discloses this subject matter. Like Murdock, the Murdock Provisional discloses a “speech recognition server” contained at a “network information source.” Murdock Provisional at 1:18-19 (Ex. 1011). The information source can provide content such as “video-on-demand” or “cable television.” Murdock Provisional at 2:12-14, 4:1-3. The network includes a back channel from which the speech recognition server receives speech control data from users, and a forward channel to send appropriate programming content as requested by the users. Murdock Provisional at 1:18-22.

215. The first element of claim 27 reads as follows:

a speech recognition system coupled to a wireline node in said network for receiving a back channel from a multiplicity of user sites coupled to said network, further comprising

216. Murdock discloses this element. As discussed above, the “speech recognition system” is the remote server computer 130, which includes speech recognition module 410. The remote server computer 130 is also a “wireline node” as that term is used in the ’196 Patent because it is a “a network node

providing video or cable television delivery to multiple users over physical wires between the node and the multiple user locations.” See Paragraphs 80-81, above. For example, the remote server computer 130 provides video or cable television delivery to multiple users. See Murdock at 3:31-36 (“Once the user is identified and the requested programming content is determined, the server computer 130 retrieves the requested program content from a program database and transmits the retrieved program content via the forward channel 132 to the local processing unit 120.”); see also 5:23-34 (Ex. 1010). The network the remote server 130 is connected to can provide cable television programming and interact with cable set top boxes. Murdock at 2:28-31 (“Examples of the requested programming content include web-based content, video-on-demand, cable television programming, and the like.”), 4:14-15 (“Examples of the local processing unit 120 may include set top boxes, cable boxes, and the like.”). Persons of ordinary skill in the art understand that cable television delivery and cable set top boxes operate on networks that are connected over physical wires, as the name “cable” suggests. Murdock thus discloses a speech recognition system that is coupled to a wireline node—the remote server computer 130.

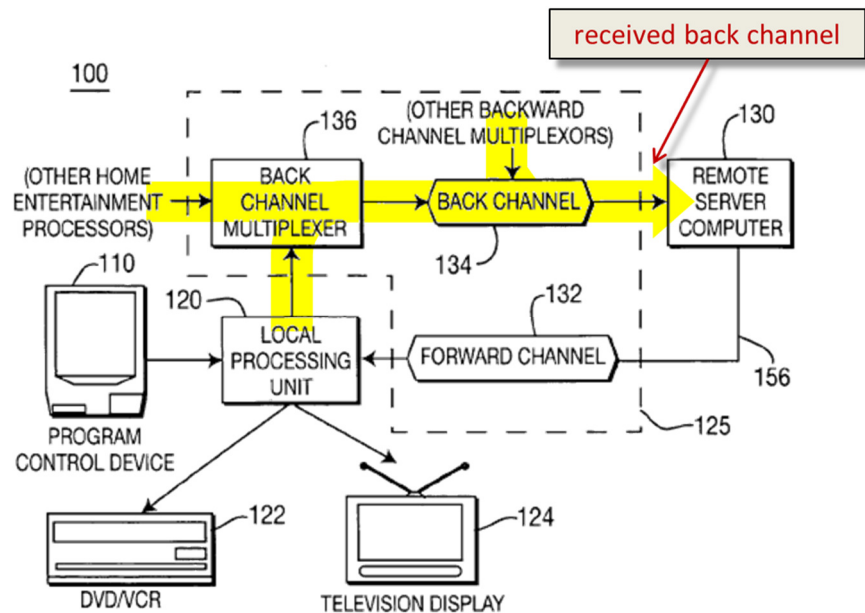
217. Murdock also discloses that the remote server computer “receiv[es] a back channel from a multiplicity of user sites coupled to said network” as this claim requires. A back channel is an “upstream communication channel delivering

signals from multiple user sites to a central wireline node.” See Paragraph 82-83.

In particular, Murdock states that the “local processing unit ... transmits the received first audio signal to a service provider via the back channel” and the “remote server computer receives the first audio signal from the back channel.”

Murdock at Abstract; see also 3:15-17 (“[T]he remote server computer 130 receives the multiplexed signal from the back channel 134 and performs speech recognition on the received signal.”), 5:2-4 (“The command signal interface 402 [of the remote server computer] receives the multiplexed signal from the back channel 134.”) (Ex. 1010). Murdock further explains that the “back channel multiplexer 136 multiplexes or combines” the signal from local processing unit 120 “with additional audio signals from the local processing unit of other users, i.e., other users of other voice activated home entertainment systems.” Murdock at 3:1-5. The back channel also “combines the multiplexed signals from other back channel multiplexers.” Murdock at 3:7-9.

218. Figure 1 of Murdock illustrates the combined signals transmitted from multiple different users over back channel 134 to remote server computer 130, which receives the back channel.



219. Murdock at Fig. 1 (Ex. 1010) (annotated). Thus, Murdock discloses that the “speech recognition system” (i.e., remote server 130) “receiv[es] a back channel ” (i.e., the voice commands of multiple users) as recited in claim 27.

220. Murdock also discloses that the back channel is from a multiplicity of user sites coupled to the network. As discussed above, the remote server computer receives audio signals from individual local processing units. Murdock at Abstract, 2:44-67, 3:15-17, 5:2-4 (Ex. 1010). The local processing unit can be a “set top terminal, a cable box, and the like.” Murdock at 2:44-47. This type of equipment is located at a particular user site. Indeed, this is how the ’196 Patent describes user sites: “Each user site contains a Set Top Box, such as STB 180,” ’196 Patent at 2:8; see also 4:18-19 (“As in FIG. 1, each user site contains a Set Top Box, such as STB 180”) (Ex. 1001). The “receiv[ed] back channel”

thus comes from “a multiplicity of user sites coupled to said network,” as recited in claim 27.

221. The Murdock Provisional also discloses this subject matter. For example, the Murdock Provisional states:

It is the second advantage of this invention to provide a mechanism, the network's backward channel, for transmitting voice control data to the network information source. The central speech recognition server is contained at the network information source. It processes speech control data arriving through the backward channel and then sends the appropriate programming modifications through the forward channel as requested by the viewer.

Murdock Provisional at 1:16-22 (Ex. 1011). The Murdock Provisional further provides that “[t]he back channel is the channel in which the Voice Control Data from the HEP is transmitted to the information source.” Murdock Provisional at 2:9-11. In other words, the information source receives the back channel containing voice control data to create a received back channel. This is illustrated in Figure 1 of the Murdock Provisional is essentially the same as Figure 1 of Murdock. It shows the Information Provider, which contains the Information Source (Murdock Provisional at 4:1-2), receiving the back channel to create a received back channel. Murdock Provisional at Fig. 1. The Information Source is “located at the site of the information provider,” which can provide “web content, video-on-demand, or cable television.” Murdock Provisional at 4:1-4. The

Murdock Provisional also discloses that individual users send speech signals from a hand-held controller and home entertainment processor to the central speech recognition server via a back channel. Murdock Provisional at 2:1-3:2.

222. The second element of claim 27 reads as follows:

a back channel receiver, for receiving said back channel to create a received back channel;

223. This element is similar to the first element of claim 1, with the exception that while claim 1 simply requires “receiving said back channel to create a received back channel,” this element requires “*a back channel receiver, for receiving said back channel to create a received back channel.*” Compare ’196 Patent at 51:1-2 with ’196 Patent at 55:15-16 (claim 27) (Ex. 1001) (emphasis added). As discussed in connection with claim 1, and in connection with the previous limitation, Murdock discloses that the remote server computer 130 receives a back channel to create a received back channel. See Paragraphs 140, 141, 217-220, above, Murdock at Abstract, 3:1-5, 3:7-9, 3:15-17, 5:2-4, Fig. 1 (Ex. 1010). The remote server computer 130 is thus the “back channel receiver” recited in claim 27.

224. The Murdock Provisional also discloses this subject matter. As also explained with respect to the first element of claim 1 and the previous limitation, the central speech recognition server at the network information source receives

speech control data sent from home entertainment processors (HEPs) via the back channel. See Paragraph 142, 221, above, Murdock Provisional at 1:16-22, 2:9-11, Fig. 1 (Ex. 1011). The central speech recognition server is thus a “back channel receiver” that “receiv[es] said back channel to create a received back channel” as recited in claim 27.

225. The third element of claim 27 reads as follows:

**a speech channel partitioner, for partitioning said received back channel into a multiplicity of received identified speech channels;
and**

226. Murdock discloses this element. This element is similar to the second element of claim 1, with the exception that while claim 1 simply requires “partitioning said received back channel into a multiplicity of received identified speech channels,” this element requires “*a speech channel partitioner, for partitioning said received back channel into a multiplicity of received identified speech channels.*” Compare ’196 Patent at 51:1-2 with ’196 Patent at 55:15-16 (claim 27) (Ex. 1001) (emphasis added). As discussed in connection with the second element of claim 1, Murdock discloses that the remote server computer 130 receives a multiplexed signal containing the combined voice signals of multiple different users over the back channel. Murdock at 3:1-12 (Ex. 1010). Murdock further discloses that once received, the remote computer server identifies and recognizes the user that generated the audio signal, as well as the programming

content requested by the user. Murdock at 3:18-31. As explained in detail with respect to claim 1, a person of ordinary skill in the art would understand that to make any sense of the multiplexed signal, such as using it to identify or recognize a specific user and to determine the programming the particular user requested, the remote server computer must include a demultiplexer to demultiplex the multiplexed signal, which constitutes “partitioning”. See Paragraphs 84-89, 145 above. Thus, for the reasons detailed in connection with the second limitation of claim 1, Murdock alone discloses a “speech channel partitioner” (remote server 130) that partitions the received back channel into a multiplicity of received identified speech channels. See Paragraphs 144-146, above. Likewise, the Murdock Provisional also discloses the same subject matter. See Paragraph 147, above.

227. As also detailed above in connection with the second element of claim 1, Nazarathy and Quigley also both disclose demultiplexing (i.e. partitioning), at a remote server (for example, a cable headend), multiplexed signals coming from multiple cable set top boxes. See Paragraphs 148-151 above. For the reasons discussed in connection with the second element of claim 1, the combinations of Murdock with Nazarathy and Murdock with Quigley also disclose this element of claim 27, where the “speech channel partitioner” would be the remote server 130. See Paragraphs 148-151 above.

228. In sum, as discussed with respect to the second element of claim 2, Murdock alone, and Murdock in combination with either Nazarathy or Quigley, disclose that the remote server 130 is a “speech channel partitioner” as recited in claim 27, and renders this element of claim 27 obvious.

229. The fourth element of claim 27 reads as follows:

a processor network executing a program system comprised of program steps residing in memory accessibly coupled to at least one computer in said processor network;

230. Murdock discloses this element. The “program steps” this element refers to are (1) “processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content;” and (2) “responding to said identified speech content to create an identified speech content response, for each of said multiplicity of said identified speech contents,” which follow immediately after this limitation. ’196 Patent at 55:24-31 (Ex. 1001). As explained below in connection with those elements, the remote server computer 130 processes the multiplicity of identified speech channels (the user’s voice commands) by identifying the user who issued the command and the programming content requested by the user. See Paragraph 237, below. Once identified, the remote server computer then responds by transmitting the requested programming content to the user. See Paragraph 239, below.

231. Murdock discloses that both of these steps are performed by the

remote server computer 130, which is shown in Figure 4 below. See Paragraphs 237, 239, below.

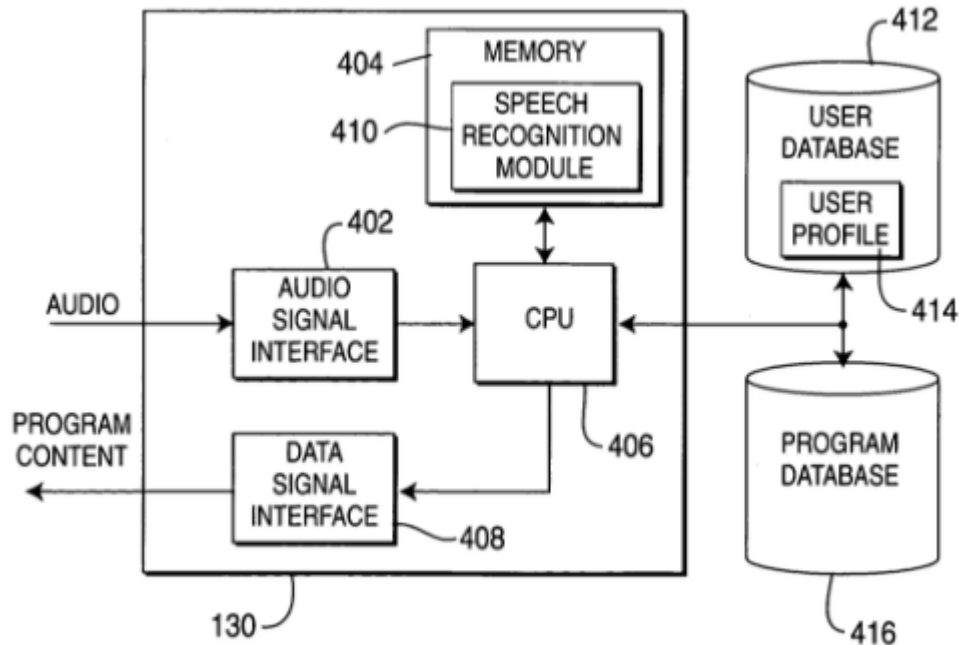


FIG. 4

Murdock Fig. 4 (Ex. 1010). As Figure 4 shows, the remote server computer 130 includes a memory 404, which “stores a speech recognition module 410 that, when retrieved and executed by the CPU 406, causes the remote server computer 130 to operate as a speech recognition server.” Murdock at 4:65-5:10. The remote server computer also “retrieves the requested programming content from the program database 416” to transmit to the user. Murdock at 5:29-31. Murdock further discloses that the remote server computer 130 can “use separate computers” to implement the functions of speech processing and transmission of programming content. Murdock at 3:38-41, 5:35-38.

232. When implemented using separate computers, a “processor network” executes the processing and responding steps, which reside in memory accessibly coupled to at least one computer, as recited in this claim. A person of ordinary skill in the art understands that a “computer,” such as a remote server computer, includes a processor that executes steps residing in memory accessibly coupled to that computer. Thus, if separate computers are used for the speech processing and transmission steps, then at least two processors, which are coupled together to perform speech recognition and transmit the requested content, are involved—one for the speech processing computer, and another for the transmission computer. The processors of these two computers form a “processing network” as recited in the claim.

233. Murdock also explicitly states that the speech recognition module 410, which allows the remote server computer to perform speech recognition on voice commands (“processing”), is stored in the remote server—i.e. coupled to the remote server—in memory 404. Murdock at 4:65-5:10 (Ex. 1010). Likewise, a person of ordinary skill in the art would understand that the instructions for transmitting the requested content to the user would also be stored in memory accessibly coupled to the computer responsible for transmitting the requested content to the user. Murdock thus discloses a “processor network executing a program system comprised of program steps residing in memory accessibly

coupled to at least one computer in said processor network” as recited in this claim.

234. The Murdock Provisional also discloses this subject matter. The Murdock Provisional discloses that the “central speech recognition server” is “contained at the network information source,” which can contain “web content, video-on-demand, or cable television.” Murdock Provisional at 1:18-19, 4:1-4 (Ex. 1011). The central speech recognition server “processes speech control data arriving through the backward channel” and then “sends the appropriate programming modifications through the forward channel as requested by the user.” Murdock Provisional at 1:18-22. Persons of ordinary skill in the art would understand that instead of using one computer to both process the speech control data and transmit the programming modifications, separate computers could be used.

235. Persons of ordinary skill in the art also understand that a server is simply a type of computer, and that for a server, such as the central speech recognition server to “process speech control data” and “send[] the appropriate programming modifications” to the viewer, the server must have program steps—*i.e.*, instructions—that reside in a memory accessibly coupled to the server.

236. The fifth element of claim 27 reads as follows:

**wherein said program system is comprised of the program steps
of:
processing said multiplicity of said received identified speech**

channels to create a multiplicity of identified speech content;

237. This element is identical to the third element of claim 1. See Paragraph 59 above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 1. See Paragraphs 153-154, above. In short, the remote server computer 130 processes a multiplicity of received identified speech channels by performing speech recognition on the verbal requests for programming content it receives from local processing units to create a multiplicity of identified speech content—the identity of the requested content. Murdock at 2:24-26, 3:1-17, 3:28-31, 5:11-14 (Ex. 1010). The Murdock Provisional also discloses these concepts. See Murdock Provisional at 1:18-22 (Ex. 1011).

238. The sixth element of claim 27 reads as follows:

responding to said identified speech content to create an identified speech content response, for each of said multiplicity of said identified speech contents; and

239. This element is nearly identical to the fourth element of claim 1. See Paragraph 59 above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 1. See Paragraphs 156-157, above. In short, after recognizing the voice commands from each user (i.e., the identified speech content for each user), the remote server computer 130

responds by retrieving the requested programming and transmitting it to the viewer. Murdock at 3:31-36, 5:23-24 (Ex. 1010). The Murdock Provisional also discloses these concepts. Murdock Provisional at 1:18-22 (Ex. 1011).

240. The last element of claim 27 reads as follows:

wherein said network supports at least one of the collection comprising: cable television delivery to said multiplicity of user sites; and video delivery to said multiplicity of user sites.

241. The Murdock system may be used to distribute “video-on-demand” and “cable television programming, and the like.” Murdock at 2:18-31, 5:26-29 (Ex. 1010). The Murdock network thus supports *both* cable television delivery *and* video delivery to a multiplicity of user sites, as recited in this element.

242. Likewise, the Murdock Provisional also discloses that the content distributed by its system may be “video-on-demand” and “cable television.” Murdock Provisional at 2:12-14, 4:1-4 (Ex. 1011).

9. Claim 28

243. Claim 28 is a dependent claim that adds the following limitations:

The system of claim 27, wherein at least one of said identified speech channels has an associated user site; and

244. Other than minor differences in their preambles (claim 2 recites a “method” whereas claim 28 recites a “system), this element is identical to the first element of claim 2. See Paragraph 60, above. Murdock and the Murdock

Provisional disclose this limitation for the reasons discussed in connection with claim 2. See Paragraphs 159-160, above. In short, the '196 Patent states that a "user site" contains a set top box. '196 Patent at 2:8, 4:18-19 (Ex. 1001).

Murdock discloses that "the identified speech channels" come from audio signals transmitted from local processing units, which can be set top boxes. Murdock at 2:44-67, 3:15-17, 5:2-4 (Ex. 1010). The Murdock Provisional also discloses this subject matter. See Paragraph 160, Murdock Provisional at 2:1-3:2 (Ex. 1011).

245. The next element of claim 28 recites:

**wherein partitioning said received back channel is further comprised of:
partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site;**

246. This limitation is substantively the same as the corresponding limitation recited in claim 2. See Paragraph 60, above. Murdock, alone or in combination with Nazarathy and Quigley, and the Murdock Provisional, disclose this limitation for the reasons discussed in connection with claim 2. See Paragraphs 162-167, above. In short, Murdock discloses that audio signals transmitted from individual local processing units are multiplexed together and sent to the remote server computer via back channel 134. Murdock at Abstract, 2:44-67, 3:15-17, 5:2-4 (Ex. 1010). Persons of ordinary skill in the art understand that the multiplexed signal received via back channel 134 would be demultiplexed

(i.e., partitioned) to identify signals from individual users. See Paragraph 145.

Thus, the “received back channel” (i.e., back channel 134) is partitioned “into a multiplicity of said received identified speech channels from said associated user site” (i.e., local processing unit 120), as recited in the claim.

247. In addition, each “received identified speech channel” from a particular “user site” may be further partitioned. For example, a user located at one user site may issue multiple commands. Each command issued is demultiplexed—partitioned—by the remote server computer. See Paragraph 163, above.

248. Finally, each individual signal from each local processing unit may itself be divided into discrete portions for transmission. For example, Quigley discloses that in systems that use TDMA, fragments of data can be transmitted in a plurality of segments. The headend then reassembles the segments to create each particular signal from each user site, as claimed. See Paragraph 164, above. The combination of Murdock and Quigley thus also render this element obvious.

249. The Murdock Provisional also discloses the same subject matter. See Paragraphs 166-167.

250. The next element of claim 28 recites:

wherein the program step processing said multiplicity of said received identified speech channels is further comprised of the program step of:

processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech content; and

251. This limitation is the same as the corresponding limitation recited in claim 2, with the exception that while claim 2 recites a “step,” this element recites a “program step.” See Paragraph 60, above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 2. See Paragraphs 169-171, above. In short, the remote server computer “processes” the “multiplicity of said received identified speech channels from said associated user site” (i.e., audible signals transmitted by local processing units) by performing speech recognition on each of these channels “to create a multiplicity of identified speech content” (i.e., the identity of the requested content). Murdock at 3:1-17, 3:28-31, 5:11-14 (Ex. 1010). This processing is performed every time a user at a user site uses a voice command to request programming content. The Murdock Provisional also discloses these concepts. See Murdock Provisional at 1:18-22 (Ex. 1011).

252. The next element of claim 28 recites:

**wherein the program step responding to said identified speech content is further comprised of the program step of:
responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.**

253. This limitation is the same as the corresponding limitation recited in claim 2, with the exception that while claim 2 recites a “step,” this element recites a “program step.” See Paragraph 60, above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 2. See Paragraphs 173-174, above. In short, the remote server computer responds to identified speech content by retrieving and transmitting the content requested by a user to that user at the user site. Murdock at 3:31-36, 5:23-34 (Ex. 1010). The Murdock Provisional also discloses the same subject matter. Murdock Provisional at 1:18-22 (Ex. 1011).

10. Claim 30

254. Claim 30 is a dependent claim that adds the following limitations:

**The system of claim 28, wherein the program step responding to said identified speech contents from said associated user site is comprised of the program steps of:
processing said identified speech content response to create said identified user site response; and sending said identified user site response to said identified user site.**

255. Other than minor differences in the preamble (claim 4 recites a “method” whereas claim 30 recites a “system”), this limitation is identical to the corresponding limitation recited in claim 4. See Paragraph 61, above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in

connection with claim 4. See Paragraphs 176-177, above. In short, Murdock discloses processing the user's voice commands (the identified speech content response) by performing speech recognition on a user's speech commands to identify, retrieve, and prepare for transmission the programming content requested by the user (the identified user site response). Murdock at 3:28-36, 5:23-24 (Ex. 1010). Murdock also discloses sending the requested programming to the user. Murdock at 3:31-36, 5:23-24. The Murdock Provisional also discloses the same subject matter. Murdock Provisional at 1:18-22, 2:5-14 (Ex. 1011).

11. Claim 31

256. Claim 31 is a dependent claim that adds the following limitations:

**The system of claim 28, wherein the program step responding to said identified speech contents from said associated user site is comprised of the program steps of:
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and**

257. The "assessing" step of claim 31 is identical to the "assessing" step of claim 5. See Paragraph 62, above. Accordingly, Murdock discloses this limitation for the same reasons. See Paragraphs 179-180, above. In short, Murdock discloses requesting content, which can include video-on-demand or pay-per-view, and then processing that request. See Murdock at 1:16-37, 2:24-31, 3:31-36 (Ex. 1010).

Persons of ordinary skill in the art understand that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. Thus, Murdock discloses “assessing said speech response” (i.e., assessing the recognized voice command) associated with a particular “user device” (i.e., program control device or local processing unit) to “create a financial consequence (i.e., purchase of a pay-per-view or video-on-demand movie), as recited in the claim. The Murdock Provisional also discloses the same subject matter. See Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

258. The last element of claim 31 is:

billing said user site based upon said financial consequence.

259. The “billing” step of claim 31 is identical to the “billing” step of claim 5. See Paragraph 62, above. Accordingly, Murdock alone and in combination with Banker or Gordon disclose this limitation for the same reasons. See Paragraphs 182-186, above. In short, Murdock discloses that users can request content, including video-on-demand or pay-per-view. Murdock at 1:16-37, 2:24-31 (Ex. 1010). A person of ordinary skill in the art would understand that ordering pay-per-view programming would result in a bill. See Paragraphs 179, 182, above. Murdock alone therefore discloses this limitation. The Murdock Provisional also discloses the same subject matter. See Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

260. In addition, Banker describes a system that allows users to purchase pay-per-view content by pressing the “BUY key” on a television remote and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). After following addition steps, the purchase is confirmed and the user is billed. Banker at 17:1-3, 7:60-63. Similarly, Gordon describes a system that allows a user to purchase an on-demand video subscription. Gordon at 9:64-10:9, Figs. 3B and 8 (Ex. 1016). If the user confirms the purchase, the user is billed for the subscription. Gordon at 2:60-63, 3:61-63. Because persons of ordinary skill in the art would be motivated to combine Murdock with Banker or Gordon (see Paragraphs 300-301 below), such combinations would also disclose this limitation.

12. Claim 32

261. Claim 32 is a dependent claim that adds the following limitations:

**The system of claim 28, wherein the program step responding to said identified speech content from said associated user site is comprised of the program steps of:
assessing said speech response to create a financial consequence identified as to said user site;**

262. The “assessing” step of claim 32 is identical to the “assessing” step of claim 6. See Paragraph 63, above. Accordingly, Murdock discloses this limitation for the same reasons. See Paragraphs 188-189, above. In short, Murdock discloses

requesting content, which can include video-on-demand or pay-per-view, and then processing that request. See Murdock at 1:16-37, 2:24-31, 3:31-36 (Ex. 1010).

Persons of ordinary skill in the art understand that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. Thus, Murdock discloses “assessing said speech response” (i.e., assessing the recognized voice command) associated with a particular “user device” (i.e., program control device or local processing unit) to “create a financial consequence (i.e., purchase of a pay-per-view or video-on-demand movie), as recited in the claim. The Murdock Provisional also discloses the same subject matter. See Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

263. In addition, both Banker and Gordon also disclose this limitation. See Paragraphs 190-191, above. For example, Banker discloses purchasing a pay-per-view movie by pushing the “BUY key” on a television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). Banker thus discloses “assessing said speech response” (i.e. assessing the user’s key entries) to “create a financial consequence” (i.e., purchase of a pay-per-view movie), as recited in the claim.

264. Likewise, Gordon discloses allowing a user to choose a movie title from a menu. Gordon at 9:18-28, Fig. 3B (Ex. 1016). After selection, the system

determines the title price. Gordon at 9:29-31, Fig. 3B. Thus, Gordon discloses “assessing said speech response” (i.e., assessing the user’s choice of a title) to “create a financial consequence” (i.e., determine the price of the title), as recited in the claim.

265. Because persons of ordinary skill in the art would be motivated to combine Murdock with Banker and Gordon (see paragraphs 300-301, below), such combinations also disclose this element.

266. The next element of claim 32 recites:

displaying said financial consequence to create a displayed financial consequence at said user site;

267. The “displaying” step of claim 32 is identical to the “displaying” step of claim 6. See Paragraph 63, above. Accordingly, Murdock discloses this limitation for the same reasons. See Paragraphs 193-194, above. In short, Murdock discloses that viewers can use voice commands to select and order pay-per-view movies. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase. The Murdock Provisional also discloses the same subject matter. See Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

268. This limitation is also disclosed in Banker and Gordon. In Banker, the

system displays a screen stating “NEXT, PLEASE PRESS [BUY] TO PURCHASE THIS EVENT” to confirm the viewer’s intent to accept the purchase (i.e., the “financial consequence”). Banker at 16:62-17:3, Fig. 6F (Ex. 1015); see Paragraph 194, above. Similarly, after the user selects a movie, Gordon queries the user to see if he wants to pay “a la carte” to view his selection. Gordon at 9:56-63; Fig. 8 (displaying price of movie) (Ex. 1016). See Paragraph 194, above.

269. Because persons of ordinary skill in the art would be motivated to combine Murdock with Banker and Gordon (see paragraphs 300-301, below), such combinations also disclose this element.

270. The next element of claim 32 recites:

confirming said displayed financial consequence from said user site to create a financial commitment; and

271. The “confirming” step of claim 32 is identical to the “confirming” step of claim 6. See Paragraph 63, above. Accordingly, Murdock discloses this limitation for the same reasons. See Paragraphs 196-197, above. In short, Murdock discloses that viewers can use voice commands to select and order pay-per-view movies. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase and confirming their intent to do so. The Murdock Provisional also discloses the same subject matter. See

Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

272. This limitation is also disclosed in Banker and Gordon. For example, to purchase a pay-per-view movie in Banker, a user must first press the BUY key, enter a system provided code, and then press the BUY key a second time to confirm the intent to purchase (i.e., the “financial consequence”) to “create a financial commitment” as claimed. Banker at 16:62-17:3, Fig. 6F (Ex. 1015); see Paragraphs 197, 198, above. Similarly, after the user selects a movie, Gordon queries the user (i.e., confirming with the user) to see if he wants to pay “a la carte” to view his selection. Gordon at 9:56-63; Fig. 8 (displaying price of movie); see also 6:43-47 (Ex. 1016); Paragraphs 197, 199, above.

273. Because persons of ordinary skill in the art would be motivated to combine Murdock with Banker and Gordon (see paragraphs 300-301, below), such combinations also disclose this element.

274. The last element of claim 32 recites:

billing said user site based upon said financial commitment.

275. The “billing” step of claim 32 is identical to the “billing” step of claim 6. See Paragraph 63, above. Accordingly, Murdock discloses this limitation for the same reasons. See Paragraphs 201, 202, above. In short, Murdock discloses that viewers can use voice commands to select and order pay-per-view movies. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010). A person of ordinary skill in the

art would have known that ordering pay-per-view programming and certain on-demand videos would require a purchase, resulting in a bill. The Murdock Provisional also discloses the same subject matter. See Murdock Provisional at 2:12-14, 2:19-33, 4:1-3 (Ex. 1011).

276. In addition, Banker describes a system that allows users to purchase pay-per-view content by pressing the “BUY key” on a television remote and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). After following addition steps, the purchase is confirmed and the user is billed. Banker at 17:1-3, 7:60-63, 16:62-17:3. Similarly, Gordon describes a system that allows a user to purchase an on-demand video subscription. Gordon at 9:64-10:9, Figs. 3B and 8 (Ex. 1016). If the user confirms the purchase, the user is billed for the subscription. Gordon at 2:60-63, 3:61-63, 10:39-47. Because persons of ordinary skill in the art would be motivated to combine Murdock with Banker or Gordon (see paragraphs 300-301 below), such combinations would also disclose this limitation.

13. Claim 38

277. Claim 38 is a dependent claim that adds the following limitation:

The system of claim 28, wherein the program step responding to said identified speech content from said associated user site is comprised, for at least one of said user sites, of the program step

of:

responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech contents identified as to said user site.

278. Other than minor differences in the preambles (claim 12 recites a “method” whereas claim 38 recites a “system” and adds the language “for at least one of said user sites”), claim 12 is identical to claim 38. See Paragraph 64. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 12. See Paragraph 204-207, above. In short, in Murdock, users request programming content by issuing spoken commands, such as “Show me all the college football games on this Saturday,” which persons of ordinary skill in the art would recognize as a “natural language” command. Murdock at 1:41-49, 2:22-26, 3:42-57 (Ex. 1010). Murdock also discloses that after identifying the programming content requested by a particular user, the server computer responds by retrieving the requested programming and transmitting it to the user. Murdock at 3:31-36, 5:23-24. The Murdock Provisional also discloses the same subject matter. Murdock Provisional at 1:18-22, 2:19-20 (Ex. 1011).

14. Claim 39

279. Claim 39 is a dependent claim that adds the following limitation:

The system of claim 28, wherein the program step processing said multiplicity of said received speech channels is further comprised,

**for at least one of said user sites, of the program step of:
processing said received speech channels from said user site based
upon natural language for said user site to create said speech
content identified as to said user site.**

280. Other than minor differences in the preamble (claim 13 recites a “method” whereas claim 39 recites a “system” and adds the language “for at least one of said user sites”), this limitation is identical to the corresponding limitation recited in claim 13. See Paragraph 65, above. Murdock and the Murdock Provisional disclose this limitation for the reasons discussed in connection with claim 13. See Paragraphs 209-211, above. In short, Murdock discloses using a natural language command, such as “Show me all the college football games on this Saturday” to request programming content. Murdock at 3:54-57 (Ex. 1010). The remote server computer processes this command by performing speech recognition on it to identify the user requesting the content and the requested content. Murdock at 3:13-17, 3:28-31, 5:11-14. The Murdock Provisional also discloses the same subject matter. Murdock Provisional at 1:18-22, 2:19-20 (Ex. 1011).

15. Claim 40

281. Claim 40 is a dependent claim that adds the following limitation:

**The system of claim 28, wherein said network includes a wired
residential broadband network servicing at least part of said**

multiplicity of user sites.

282. A person of ordinary skill in the art understands a “broadband network” to refer to a network capable of transmitting data at high speeds, such as networks made up of coaxial or fiber-optic cables. This term originated to contrast connectivity means from slower dialup network access using conventional telephone modems. This understanding is also reflected in the ’196 Patent. For example, the ’196 Patent states that “FIG. 2 depicts a typical residential *broadband* network using local loop wiring of the network, as disclosed in the prior art.” ’196 Patent at 4:15-17 (Ex. 1001) (emphasis added). A few lines earlier, the ’196 Patent named Hybrid-Fiber-Copper (HFCop) and hybrid fiber-coaxial (HFC) networks as examples of local loop networks. ’196 Patent at 3:63-4:9. Thus, even according to the ’196 Patent, networks made up of fiber-optic, coaxial, and even copper wires, are broadband networks.

283. The network disclosed in Murdock supports cable television. Murdock at 2:30 (requested programming can include “cable television programming”), 2:44-47 (examples of local processing units include “cable boxes” and the like) , 4:13-15 (same) (Ex. 1010). Persons of ordinary skill in the art at the time of the ’196 Patent understood that cable networks were typically made up of coaxial lines or hybrid-fiber-coaxial networks (“HFC” networks). ’196 Patent at 2:51-54 (Ex. 1001); Quigley at 10:45-48 (Ex. 1014). These networks provided

sufficient bandwidth for the transmission of data at high speeds. The network in Murdock is therefore wired and broadband. Moreover, as explained above, the network connects to local processing units, which are located at user sites.

Murdock at 2:44-47; 4:13-15. The network is therefore also “residential” and services “at least part of said multiplicity of user sites” as required by the claim.

Murdock alone thus discloses this claim.

284. The Murdock Provisional also discloses the same subject matter. For example, it also discloses that its network can be used to transmit cable television. Murdock Provisional at 4:3 (Ex. 1011).

285. In addition, both Quigley and Nazarathy also disclose HFC cable networks (i.e. broadband networks). For example, Quigley’s system is built for an HFC network. Quigley at 8:55-60, 10:41-11:3 (Ex. 1014). Likewise, the entire disclosure of Nazarathy is directed toward a hybrid fiber coaxial cable network. See, e.g., Nazarathy at Abstract, 9:25-26, (“This invention pertains to broadband HFC distribution systems to digital home terminals”) (Ex. 1013). It would have been natural for a person of ordinary skill in the art to implement Murdock in an HFC network as disclosed by Quigley and Nazarathy, given the prevalence of HFC networks used for cable television at the time of the ’196 Patent. Persons of ordinary skill in the art would also have been capable of implementing the Murdock voice control system in an HFC network, as it would have been no more

than combining prior art elements (the Murdock system and HFC networks) to yield a predictable result (the Murdock system running on an HFC network). Moreover, as discussed below, persons of ordinary skill in the art would have been motivated to combine Murdock with Quigley and Nazarathy. See Paragraphs 294-299. Thus, the combination of Murdock with Quigley or Nazarathy also disclose this element.

16. Claim 41

286. Claim 41 is a dependent claim that adds the following limitation:

The system of claim 40, wherein said wired residential broadband network is a fiber-to-the-curb residential broadband network.

287. Murdock discloses this element. As discussed with respect to claim 40, Murdock discloses a “wired residential broadband network” that is “residential.” See Paragraphs 282-285 above. This limitation additionally requires that the network be “fiber-to-the-curb.” ’196 Patent at 57:33-34 (Ex. 1001). The ’196 Patent does not define the phrase “fiber-to-the-curb”; indeed, the only time this phrase is used is in claim 41 of the patent. Without further guidance, a person of ordinary skill in the art would understand a “fiber-to-the-curb” network to mean a network that runs fiber-optic cables from a central server or headend to a fiber termination node which serves multiple users in the neighborhood of the fiber termination. Other types of cables, such as coaxial cables or copper wires (twisted

pair) connect the multiple user sites to the node. Thus, one example of a fiber-to-the-curb network is a hybrid-fiber-coaxial (HFC) network. In HFC networks, optical fibers run from a central office, or headend, to a neighborhood node. '196 Patent at 3:44-45 (Ex. 1001). "At the node, coaxial cable connects" multiple users to the node. '196 Patent at 3:49-52; see also Quigley at 8:55-60 (defining HFC as networks "wherein optical fiber provides communication to a plurality of fiber nodes and each fiber node typically serves approximately 500 to 2,000 subscribers, which communicate with the node via coaxial cable.") (Ex. 1014).

288. The network disclosed in Murdock supports cable television. Murdock at 2:30 (requested programming can include "cable television programming"), 2:44-47 (examples of local processing units include "cable boxes" and the like) , 4:13-15 (same) (Ex. 1010). Persons of ordinary skill in the art at the time of the '196 Patent understood that cable networks were typically made up of coaxial lines or hybrid-fiber-coaxial networks ("HFC" networks). Indeed, the '196 Patent recognized that, at the time of its filing, "[m]ost current cable systems use the popular HFC architecture." '196 Patent at 2:51-54 (Ex. 1001). Other prior art references also demonstrate that HFC networks were common and well known at the time of the '196 Patent. For example, Quigley states that "hybrid fiber coaxial networks are commonly utilized by cable providers to provide Internet access, cable television, pay-per-view and the like to subscribers." Quigley at 10:45-48

(Ex. 1014). Therefore, it would have been natural for a person of ordinary skill in the art reading Murdock to understand that its network can be an HFC network—a “fiber-to-the-curb” network. Thus, Murdock alone discloses this claim.

289. The Murdock Provisional also discloses the same subject matter. For example, it also discloses that its network can be used to transmit cable television. Murdock Provisional at 4:3 (Ex. 1011).

290. Moreover, both Quigley and Nazarathy also disclose HFC cable networks. For example, Quigley’s system is built for an HFC network. Quigley at 8:55-60, 10:41-11:3 (Ex. 1014). Likewise, the entire disclosure of Nazarathy is directed toward a hybrid fiber coaxial cable network. See, e.g., Nazarathy at Abstract, 9:25-26, (“This invention pertains to broadband HFC distribution systems to digital home terminals”) (Ex. 1013). It would have been natural for a person of ordinary skill in the art to implement Murdock in an HFC network as disclosed by Quigley and Nazarathy, given the prevalence of HFC networks used for cable television at the time of the ’196 Patent. As discussed below, persons of ordinary skill in the art would have been motivated to combine Murdock with Quigley and Nazarathy. See Paragraphs 294-299. Persons of ordinary skill in the art would also have been capable of implementing the Murdock voice control system in an HFC network, as it would have been no more than combining prior art elements (the Murdock system and HFC networks) to yield a predictable result

(the Murdock system running on an HFC network). The combination of Murdock with the HFC networks disclosed in Quigley or Nazarathy thus also disclose claim 41.

17. Claim 42

291. Claim 42 is a dependent claim that adds the following limitation:

The system of claim 40, wherein said wired residential broadband network is a fiber-to-the-home residential broadband network.

292. Murdock discloses this limitation. As discussed with respect to claim 40, Murdock discloses a “wired residential broadband network” that is “residential.” See Paragraphs 282-285 above. This limitation additionally requires that the network be “fiber-to-the-home.” ’196 Patent at 57:33-34 (Ex. 1001). The ’196 Patent does not define the phrase “fiber-to-the-home”; indeed, the only time this phrase is used is in claim 42 of the patent. Without further guidance, a person of ordinary skill in the art would understand a “fiber-to-the-home” network to mean a network that runs fiber-optic cables from a central server or headend to a termination at a home, such as a user site. Other cables, such as coaxial cables, may then be used to distribute the signal from the home termination point to other points in the home itself. In other words, a fiber-to-the-home network is simply another implementation of an HFC network, where the fiber termination point ends at individual homes, as opposed to a neighborhood node in a fiber-to-the-curb

network.

293. With this understanding, Murdock alone and in combination with Nazarathy or Quigley discloses this element. As discussed in connection with claim 41, a person of ordinary skill in the art would understand that Murdock can be implemented in an HFC network. See Paragraphs 287-290. Persons of ordinary skill in the art would also have combined the system of Murdock with the HFC networks described in Nazarathy and Quigley. See Paragraphs 294-299. In addition, persons of ordinary skill in the art also knew at the time of the '196 Patent that the HFC network used could be a fiber-to-the-home network, as such systems were well-known at the time of the '196 Patent. In fact, fiber-to-the-home networks were experimentally deployed since the mid-1980s. I personally saw the Hi-Ovis fiber-to-the-home system in Japan in approximately 1986. Persons of ordinary skill in the art would also have been capable of implementing the Murdock voice control system in an HFC network, as it would have been no more than combining prior art elements (the Murdock system and HFC networks) to yield a predictable result (the Murdock system running on an HFC network). Murdock, alone or in combination with Nazarathy or Quigley, thus also disclose this element. The Murdock Provisional also discloses the same subject matter. For example, it also discloses that its network can be used to transmit cable television. Murdock Provisional at 4:3 (Ex. 1011).

B. A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Murdock with the Teachings of Nazarathy and Quigley

294. A person of ordinary skill in the art would have been motivated to combine Murdock with the teachings of Nazarathy and Quigley. Murdock, Nazarathy, and Quigley all disclose inventions for use in two-way transmission of data in cable networks. For example, Murdock states that users can request programming such as “video-on-demand” and “cable television programming.” Murdock at 2:28-31 (Ex. 1010). Murdock also discloses that its local processing unit may be a “set top terminal, a cable box, and the like”—components of a cable television system. Murdock at 2:44-47. Similarly, Nazarathy discloses a “distributed termination system” for use in “hybrid fiber coaxial cable networks such as used in cable television where two-way digital communications are desired.” Nazarathy at Abstract (Ex. 1013). In Nazarathy’s system, the “home terminals” can be “digital set-top boxes or cable models using bi-directional transmission standards.” Nazarathy at Abstract. Likewise, Quigley discloses a “cable modem apparatus and method” that describes a “number of features for enhancing the performance of a cable transmission system.” Quigley at Abstract (Ex. 1014). Quigley also describes the cable network it discloses as consisting of a “hybrid fiber coaxial (HFC) network, wherein optical fiber provides communication to a plurality of fiber nodes,” where a “plurality of subscribers

communicate with the fiber node via a common or shared coaxial cable.” Quigley at 8:55-62. In short, the networked systems of Nazarathy and Quigley are designed to be used in the very same network infrastructure disclosed by Murdock—cable networks.

295. Further, Murdock, Nazarathy, and Quigley also describe systems requiring bi-directional communication between a central node (like a headend or a remote server) and multiple, distributed user units. Murdock, for example, describes a system where audio signals from multiple users are transmitted through the users’ local processing units to a central remote server computer 130, which then processes the audio signals and responds by providing the users with requested programming content. Murdock at 2:11-3:41, see also 1:49-55 (“The local processing unit...transmits the received first audio signal via the back channel. The remote server computer...transmits the programming content to the local processing unit via the forward channel.”) (Ex. 1010). Nazarathy likewise states that its cable modem termination system handle “the upstream digital traffic arriving over digital fibers from multiple nodes into a single digital stream aggregating the individual digital upstream transmissions from the nodes.” Nazarathy at 14:50-54, Figure 9, see also 13:7-9 (“The MAC domain consists of the set of all users served by one or more downstream channels and one or more upstream channels, bound together by the two-way MAC protocol.”) (Ex. 1013).

Similarly, Quigley states that its cable modem termination system “must be capable of facilitating bidirectional communication with any desired one of the plurality of cable modems,” and provides details for how this bidirectional communication may happen. Quigley at 8:46-48, see also 9:18-25 (in the “Detailed Description of the Invention,” noting a “downstream” data rate of up to 56 Mbps and an “upstream” data rate of up to 10 Mbps) (Ex. 1014).

296. Murdock, Nazarathy, and Quigley are also similar in that all three describe transmitting media content in the downstream direction. Murdock at 2:28-30, 3:31-36, 5:23-34 (Ex. 1010); Nazarathy at 1:15-44, 1:65-2:7 (Ex. 1013); Quigley at 11:4-7, 13:11-21, 16:31-43 (Ex. 1014). Murdock and Quigley further describe systems that receive voice commands in the upstream (from user to headend) direction. In Murdock, users provide “input verbal command signal[s],” such as requests for video-on-demand or cable television programming, that are transmitted to the remote server computer 130, which then retrieves and transmits the requested programming back to the requesting user. Murdock at 2:22-3:17, 3:31-36. Quigley also states that upstream communications can include “voice,” and that its headend, which contains the CMTS, may include “video servers.” Quigley at 11:4, 63:44-46.

297. Nazarathy and Quigley also provide additional details for implementing bi-directional communication systems for cable networks that fit

directly into, but are not expressly described, by Murdock. For example, Murdock discloses that voice commands from multiple users are multiplexed together for transmission to the remote server computer, but does not otherwise provide other details regarding how this can be done. Murdock also does not expressly describe a physical network architecture for its system. Nazarathy provides specific examples of multiplexing, which can be implemented in Murdock, such as multiplexing signals in the “return path” using, for example, time division multiplexing (TDM) and wavelength division multiplexing (WDM) for transmission to the cable headend. Nazarathy at Abstract, 13:64-14:46 (Ex. 1013). Nazarathy explicitly teaches that these multiplexed signals are demultiplexed at the cable headend. Nazarathy at 14:62-64 (“Any operations of TDM and/or WDM multiplexing are undone at the HE [cable headend] by corresponding WDM and TDM demultiplexers.”). Nazarathy further teaches that bi-directional cable networks can be implemented as hybrid-fiber-coaxial (HFC) networks. Nazarathy at 1:6-9. Quigley also describes how the “upstream channel” of the cable network can be “partitioned” into a plurality of time intervals. Quigley at 46:32-40 (Ex. 1014). This partitioning is necessary to use “TDMA for upstream communication.” Quigley at 9:56-10:3. Quigley further describes demultiplexing of multiplexed signals. Quigley at 51:49-53. Like Nazarathy, Quigley also explains that HFC networks can be used for cable television networks. Quigley at

8:55-60, 10:41-48. Thus, Nazarathy and Quigley disclose details for multiplexing / demultiplexing signals and cable television network architecture not expressly described by Murdock. A person of ordinary skill in the art, looking for further details regarding how to implement Murdock in a cable system, would thus naturally consider the teachings of Nazarathy and Quigley because they disclose methods for bi-directional communication that fit directly into the communication needs of Murdock, and would know that the system of Murdock could be implemented using the techniques and networks described in Nazarathy and Quigley.

298. A person of ordinary skill in the art would have recognized the benefits of combining Murdock with the teachings of Nazarathy and Quigley. Nazarathy is directed at providing more upstream bandwidth per user by multiplexing signals at intermediate network “hubs” in a hybrid fiber coaxial (HFC) network. Nazarathy at 3:37-43 (Ex. 1013). Quigley is directed to avoiding any undesirable reductions in upstream data rates resulting from a plurality of subscribers competing for bandwidth over a common coaxial cable (i.e., collision avoidance). Quigley at 1:46-58 (Ex. 1014) In a combined system, the upstream architectures for cable television systems disclosed in Nazarathy and Quigley would be combined with Murdock’s speech recognition system, providing more upstream bandwidth per user, allowing for higher fidelity audio commands and a

better rate of recognition. The combined system would also improve the user interfaces of the Nazarathy and Quigley systems, as the addition of Murdock's voice control system would allow users to more easily search for, request, and obtain desired programming content. Moreover, a person of ordinary skill in the art would have readily been able to combine the upstream architecture teachings of Nazarathy to Murdock's speech recognition system, to achieve predictable results, without undue effort or experimentation.

299. Because Nazarathy and Quigley describe the same general system architecture as Murdock (bi-directional communication between multiple network users on one end and a server/headend on the other end), and all three relate to sending data in the upstream direction, and media content in the downstream direction, a person of ordinary skill in the art would combine the teachings of Nazarathy and Quigley with Murdock as discussed above. This would have been no more than combining prior art elements according to known methods to yield predictable results.

C. A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Murdock with the Teachings of Either Banker or Gordon

300. A person of ordinary skill in the art would have been motivated to combine the teachings of either Banker or Gordon in the Murdock system to allow users to order pay-per-view content and movies-on-demand by voice. All three

patents (Murdock, Banker, and Gordon) relate to user interfaces for control of television systems. Murdock at 2:24-31, see also 1:16-37 (Ex. 1010); Banker at Abstract (“Apparatus for providing a user friendly interface to a subscription television terminal comprises a key pad arranged into a plurality of key groupings and an on-screen display controller for generating a plurality of screens for display on an associated television receiver.”) (Ex. 1015); Gordon at 2:41-63 (“The invention is embodied in a combination of software, which provides a so-called “navigator”, and hardware....The graphical user interface functionality is distributed between the service provider equipment and subscriber equipment....Using the inventive equipment, a consumer is capable of further subscribing to packages of on-demand programming, i.e., SOD services, through manipulation of the graphical user interface...As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that package and is billed accordingly.”); see also 1:8-14 (Ex. 1016). All three patents explicitly discuss pay-per-view functionality in their respective television systems. Murdock at 2:24-31 (“The input command signal may comprise a verbal request of programming content from a service provider....Examples of the requested programming content include web-based content, video-on-demand, cable television programming, and the like.”); see also 1:16-37 (discussing the need in the art “for a system and a concomitant method for economically provided video

content in response to a spoken command” to address pay-per-view scenarios, among others); Banker at 16:51-55 (“When the terminal of FIG. 3 is in the On mode and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by actuating the BUY key initiates a buy sequence. Consequently, this key is used to purchase an event.”) (Ex. 1015); Gordon at 1:8-13 (“The present invention relates to an interactive information distribution system such as a video-on-demand (VOD) system.”) (Ex. 1016).

301. A person of ordinary skill in the art would have recognized the benefit of combining Murdock with the teachings of either Banker or Gordon that pertain to known features and functionality of pay-per-view systems including, for example, requiring entry of a secret code and confirming user purchases. Murdock does not provide any details regarding how pay-per-view can be implemented in its system. Banker and Gordon disclose user interfaces for implementing pay-per-view functionality. Banker at 16:43-50, 16:51-55, 16:62-17:3, Figs. 6F (Ex. 1015); Gordon at 2:60-63, 3:61-63, 8:66-10:9, 10:43-46, Fig. 3B & 8 (Ex. 1016). Persons of ordinary skill in the art would therefore have naturally considered Banker or Gordon in implementing pay-per-view and video-on-demand in Murdock. Moreover, the Banker and Gordon approaches reduce the chance of an accidental or unauthorized purchase. It would also have been within the capability of a person of ordinary skill in the art to combine Murdock and either Banker or

Gordon to permit a user to order pay-per-view movies using the sequence described by Banker or Gordon but using voice commands as in Murdock. The combination would have been no more than combining prior art elements according to known methods to yield predictable results.

IX. THE CHALLENGED CLAIMS ARE RENDERED OBVIOUS BY JULIA IN COMBINATION WITH NAZARATHY, QUIGLEY, BANKER, OR GORDON.

302. It is my opinion that the challenged claims of the '196 Patent are rendered obvious to a person of ordinary skill in the art based on Julia in combination with Nazarathy, Quigley, Banker, and Gordon. Julia discloses a system for accessing programming content, such as video-on-demand, using natural language voice requests that operates in the manner described by the '196 Patent. In Julia, users request programming by issuing a voice command using a remote that is paired with a communications box, such as a set top box. That request is transmitted from the communications box via a network, such as a cable television network, to a central server, which processes the signal by performing speech recognition. After identifying the requested content, the central server transmits the requested content back to the requesting viewer. Quigley explicitly states that its system can support simultaneous voice requests from multiple users, but does not provide any details regarding how such a system could be implemented. Rather, it simply assumes that persons of ordinary skill in the art

would know how to implement a system that supports multiple simultaneous requests.

303. Nazarathy and Quigley disclose combining multiple signals from multiple locales into a single signal for transmission to a server, and then partitioning that combined signal into its constituent parts. The combination of Julia with Nazarathy and Quigley thus renders the challenged claims obvious.

304. Claims 5, 6, 31, and 32 of the '196 Patent additionally require generating “financial consequences” based on the content requested by the user. Julia discloses implementing its system on video-on-demand, and Banker and Gordon provide further details regarding how a type of video-on-demand system—pay-per-view—works. Thus, Julia in combination with Banker and Gordon also render claims 5, 6, 31, and 32 obvious.

305. In the following paragraphs, I cite exemplary portions of the prior art references to show that each element of the challenged claims is disclosed by (i) Julia combined with Nazarathy; (ii) Julia combined with Quigley; (iii) Julia combined with Banker; and (iv) Julia combined with Gordon. In the section that follows, I discuss further the reasons why a person of ordinary skill in the art would have been motivated to combine Julia with the teaching of Nazarathy, Quigley, Banker, and Gordon.

A. Every Element of the Challenged Claims Is Disclosed by Julia Combined with Nazarathy, Quigley, Banker, or Gordon

1. Claim 1

306. The preamble of claim 1 reads as follows:

A method of using a back channel containing a multiplicity of identified speech channels from a multiplicity of user sites presented to a speech processing system at a wireline node in a network supporting at least one of cable television delivery and video delivery, comprising the steps of:

307. Julia discloses the method recited in claim 1. Julia uses a “back channel containing a multiplicity of identified speech channels from a multiplicity of user sites” to deliver voice commands to a remote speech processing system at a wireline node, as recited in claim 1. For example, in Julia, a “user’s voice input data is captured by a voice input device 102, such as a microphone.” Julia at 3:39-41 (Ex. 1012). This voice data, which can include requests for programming, such as “Show me the movie ‘Unforgiven’,” is “transmitted from device 102 . . . to communications box 104 (e.g., a set-top box or a similar communications device that is capable of retransmitting the raw voice data and/or processing the voice data) local to the user’s environment and coupled to communications network 106.” Julia at 3:48-54; 10:12-13; 11:31-32. “Network 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable,

traditional copper wire (twisted pair), or any other type of hardwired connection.”

Julia at 4:31-35. Julia makes clear that network server 108 can be connected to a multiplicity of client devices, each of which can issue voice data. Julia at 6:12-26.

“The voice data is then transmitted across network 106 to a remote server or servers 108.” Julia at 3:54-55. This system is illustrated in Figure 1 of Julia.

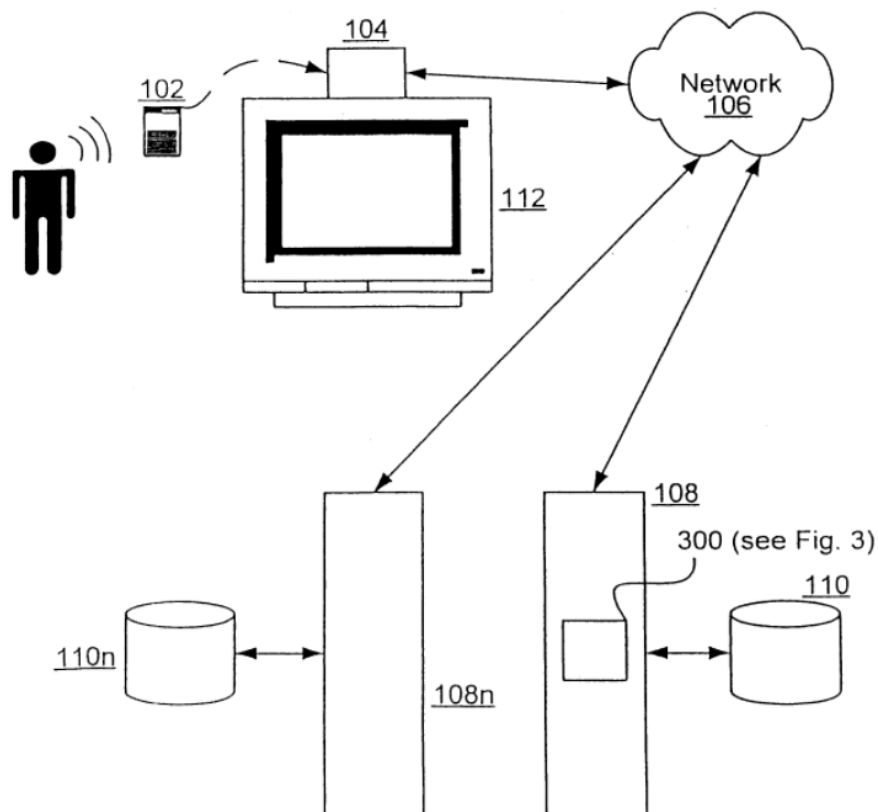


Fig. 1a

Julia at Fig. 1a. The upstream channel of network 106 is thus a “back channel” as that term is used in the ’196 Patent, because it is an “upstream communication channel delivering signals from multiple user sites to a central wireline node.” See

Paragraphs 82-83. As explained below, the wireline node is the remote server 108.

308. After receiving the voice data, the remote server “processe[s]” the voice data by performing speech recognition on it—the remote server uses “request processing logic 300 in order to understand the user's request and construct an appropriate query or request for navigation of remote data source 110.” Julia at 3:61-64 (Ex. 1012). After the desired information and programming content has been retrieved from the data source, it is transmitted from the network server 108 via the network 106 back to the client display device 112. Julia at 4:18-20, 11:60-67.

309. In Julia, the network server 108 is a “wireline node” because it is “a network node providing video or cable television delivery to multiple users over physical wires between the node and the multiple user locations.” See Paragraphs 80-81 above. For example, the network server 108 delivers videos and cable television to users over network 106. Julia at 11:60-65 (example of delivering the movie *Unforgiven* to a user); 4:31-35 (network of Julia can be a cable television network) (Ex. 1012). As shown in Fig. 1a above, network 106 connects the network server 108 with the client display devices 112, and may be embodied in cable television lines, fiber-optic cable, or other types of “hardwired” (i.e., physical) connections. Julia at 4:31-35. The network server 108 is thus a “wireline node,” as that term is used in the ’196 Patent, that is “in a network

supporting at least one of cable television delivery and video delivery,” as recited in the claim.

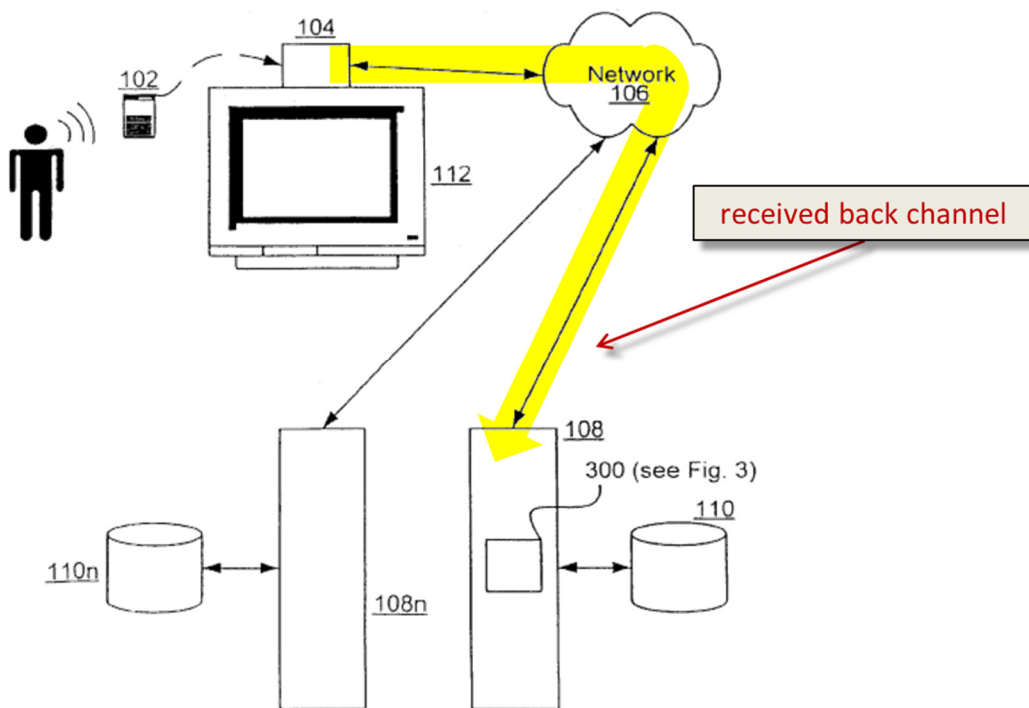
310. The first step of claim 1 of the '196 Patent is:

receiving said back channel to create a received back channel;

311. Julia discloses this element. As explained above, the “back channel” is the upstream communication channel of network 106. Julia discloses that a user’s voice input data is transmitted from the user’s “communication box 104 (e.g., a set-top box or a similar communications device that is capable of retransmitting the raw voice data and/or processing the voice data)” through network 106 to remote server 108, where the voice data is processed. Julia at 3:48-64, Fig. 1 (Ex. 1012). “Network 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.” Julia at 4:31-35. Julia also discloses that multiple network users may issue requests “simultaneously or otherwise,” and that the data source 110 and processing logic are “preferably designed and implemented to support queuing and multi-tasking of requests from multiple simultaneous network users, as will be appreciated by those of skill in the art.” Julia at 6:12-26. In other words, multiple users may send voice input data simultaneously through network 106 to remote server 108. The network server

108 therefore “receives” the back channel to “create a received back channel” (the data coming in from the multiple network users from via network 106).

312. Figure 1 of Julia illustrates the signals transmitted from a user over the back channel of network 106 to remote server 108, which receives the back channel.



Julia at Fig. 1a (Ex. 1012) (annotated). Thus, Julia discloses remote server 108 receiving a “back channel to create a received back channel,” as recited in claim 1.

313. The next step in claim 1 of the '196 Patent is:

partitioning said received back channel into a multiplicity of received identified speech channels;

314. As discussed above, the “received back channel” is comprised of

voice requests from multiple network users. Julia states that multiple network users may issue voice requests “simultaneously or otherwise,” and that the data source 110 and processing logic are “preferably designed and implemented to support queuing and multi-tasking of requests from multiple simultaneous network users, as will be appreciated by those of skill in the art.” Julia at 6:12-26 (Ex. 1012). Julia does not, however, provide any details regarding how the system can be designed to support such queuing and multi-tasking of requests from multiple simultaneous network users.

315. As “known” and “appreciated by those of skill in the art” (Julia at 6:12-26, Ex. 1012) at the time of the ’196 Patent, one way to support the simultaneous transmission and processing of voice requests from multiple network users is to multiplex the voice requests together for transmission over the back channel, and then demultiplex (partition) the voice requests at the remote server or headend.

316. The combination of Julia and Nazarathy discloses this partitioning of a received back channel as recited in claim 1 of the ’196 Patent. In particular, Nazarathy discloses a “hybrid coaxial cable network” in which data signals in the “return path” (i.e., the back channel) are multiplexed for transmission to the headend. Nazarathy at Abstract, 13:64-14:46 (Ex. 1013). For instance, Nazarathy discloses “TDM multiplexing” (i.e., time division multiplexing) and “WDM

multiplexing” (i.e., wavelength division multiplexing) the signals on the return path. Nazarathy at 14:6-8 (“There are several possibilities, e.g., WDM multiplexing, TDM multiplexing, or a combination of both WDM and TDM multiplexing”); see also 14:9-61. Julia in combination with Nazarathy thus teaches multiplexing the received back channel for transmission to the remote server.

317. Nazarathy then expressly discloses that the multiplexed signal is demultiplexed at the cable headend: “Any operations of TDM and/or WDM multiplexing are undone at the HE [cable headend] by corresponding WDM and TDM demultiplexers.” Nazarathy at 14:62-64 (Ex. 1013). Thus, Nazarathy teaches that multiplexed back channel of Julia would be demultiplexed at the remote server to identify the incoming data signals. The resulting demultiplexed signals, which constitute the voice requests of each individual user, are the “identified speech channels” recited in this element of claim 1.

318. As discussed in more detail below, a person of ordinary skill in the art would have been motivated to combine Julia and Nazarathy in this manner. See Paragraphs 440-445, below.

319. Further, Julia can also be combined with Quigley to disclose “partitioning” a received back channel as recited in claim 1 of the ’196 Patent. In particular, Quigley describes a cable network in which the “upstream channel” is

“partitioned” by dividing the channel into a plurality of time intervals. Quigley at 46:32-40 (Ex. 1014). This partitioning into “time slots” is necessary to “accomplish TDMA for upstream communication,” as disclosed in Quigley. Quigley at 9:56-10:3 (“In order to accomplish TDMA for upstream communication, it is necessary to assign time slots within which cable modems having a message to send to the cable modem termination system are allowed to transmit.”). As disclosed in Quigley, the cable modem termination system located at the headend partitions the upstream channel into time slots associated with each of the cable modems. Quigley at 9:56-10:3, Thus, the combination of Julia with the TDMA upstream communication technique of Quigley discloses “partitioning said received back channel into a multiplicity of received identified speech channels,” as recited in claim 1. Specifically, the demultiplexed signals, which constitute the voice requests of each individual user, are the “identified speech channels” recited in this element of claim 1.

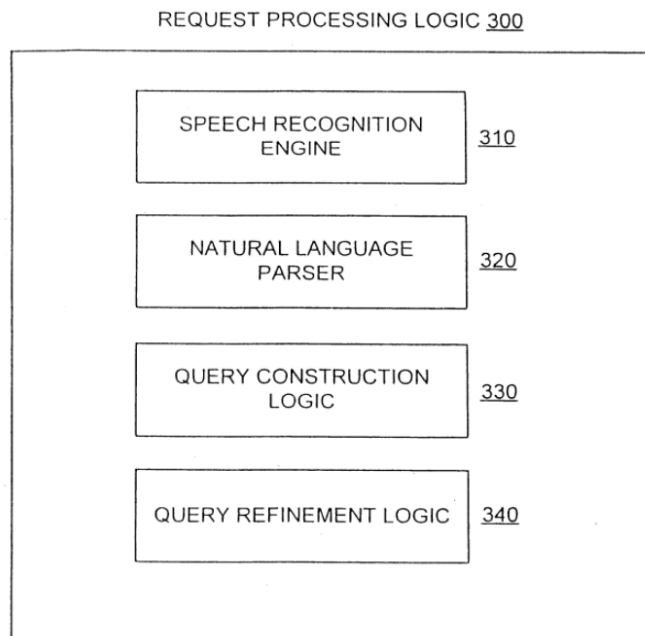
320. As discussed in more detail below, a person of ordinary skill in the art would have been motivated to employ the TDMA upstream communication technique of Quigley in the “voice activated system” of Julia. See Paragraphs 440-445, below.

321. Next, claim 1 of the ’196 Patent requires:

processing said multiplicity of said received identified speech

channels to create a multiplicity of identified speech content; and

322. Julia discloses this element. After receiving the identified speech channels—*i.e.*, the voice input data from the various users—the remote server 108 “process[es]” the voice data using “processing logic 300 in order to understand the user’s request and construct an appropriate query or request for navigation of remote data source 110, in accordance with the interpretation process exemplified in FIG. 4 and FIG. 5.” Julia at 3:61-65 (Ex. 1012). Processing logic 300 is schematically shown by Figure 3 of Julia:



Julia Fig. 3

323. Figures 4 and 5 of Julia (Ex. 1012) provide an overview of how user commands are processed in Julia:

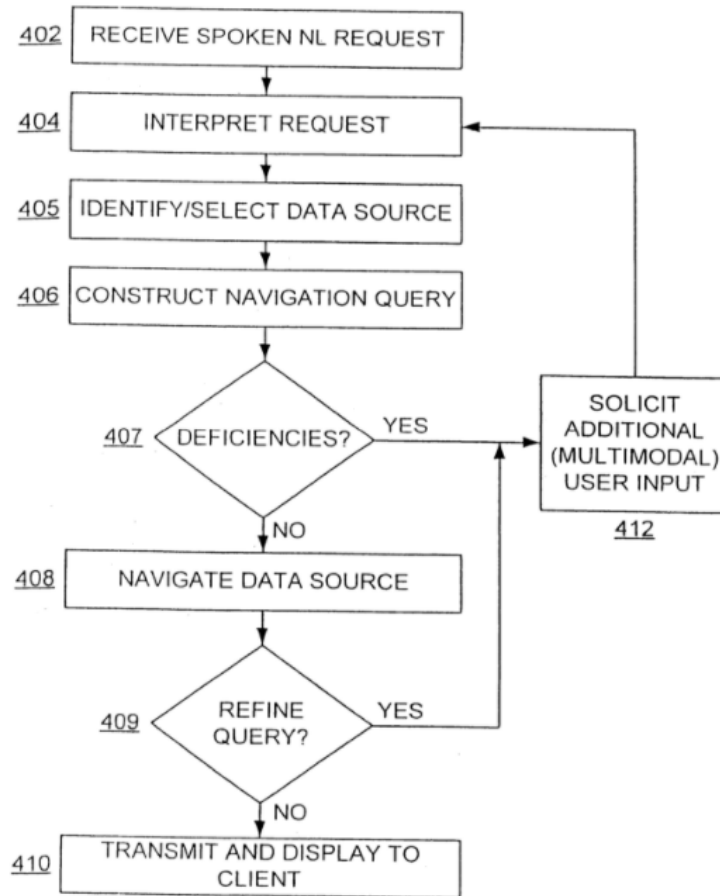


Fig. 4

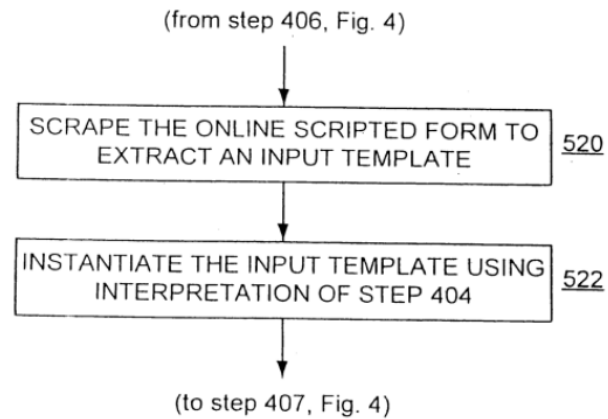


Fig. 5

324. First, “[a]t step 404, the voice data received from the user is

interpreted in order to understand the user's request for information." Julia at 7:5-7 (Ex. 1012). This can be done, for example, by first "performing speech recognition in order to extract words from the voice data, and further includes natural language parsing of those words in order to generate a structured linguistic representation of the user's request." Julia at 7:7-11. The speech recognition step 404 is done by the speech recognition engine 310, which can be, for example, commercially available engines from Nuance Communications and IBM. Julia at 7:12-46.

325. Once the words of the voice data received from the user are recognized, "the natural language interpreter attempts to determine both the meaning of the spoken words (semantic processing) as well as the grammar of the statement (syntactic processing)." Julia at 7:47-53 (Ex. 1012). One example of a natural language interpreter is the Gemini Natural Language Understanding System. Julia at 7:54; see also Julia at 7:47-8:32 (generally describing how natural language interpreters work).

326. Based on the understanding of the user's voice data generated by the natural language interpreter 320, processing logic 300 first "identifies and selects an appropriate online data source where the desired information . . . can be found," and then constructs a "navigation query" that reflects the "user's request as interpreted by the speech recognition engine and the NL parser in step 404" to

navigate the selected data source” Julia at 8:34-9:18 (Ex. 1012). Navigation queries can be “electronic query, form, or series of menu selections, or the like.” Julia at 8:48-55. Julia provides an overview of how navigation queries can be formed. Julia at 8:45-9:55.

327. Once the navigation query is formed, the identified data source is navigated and a response is generated to the user. For example, if the user’s voice command is “Show me the movie ‘Unforgiven,’” the system in Julia runs through steps listed in Figures 4 and 5, “retrieves the desired film” and then “electronically transmit[s]” it “in step 410 from network server 108 to client display device 112 via communications network 106.” Julia at 10:12-13, 11:47-67; see also 2:27-41, 2:61-67, 4:10-30 (Ex. 1012).

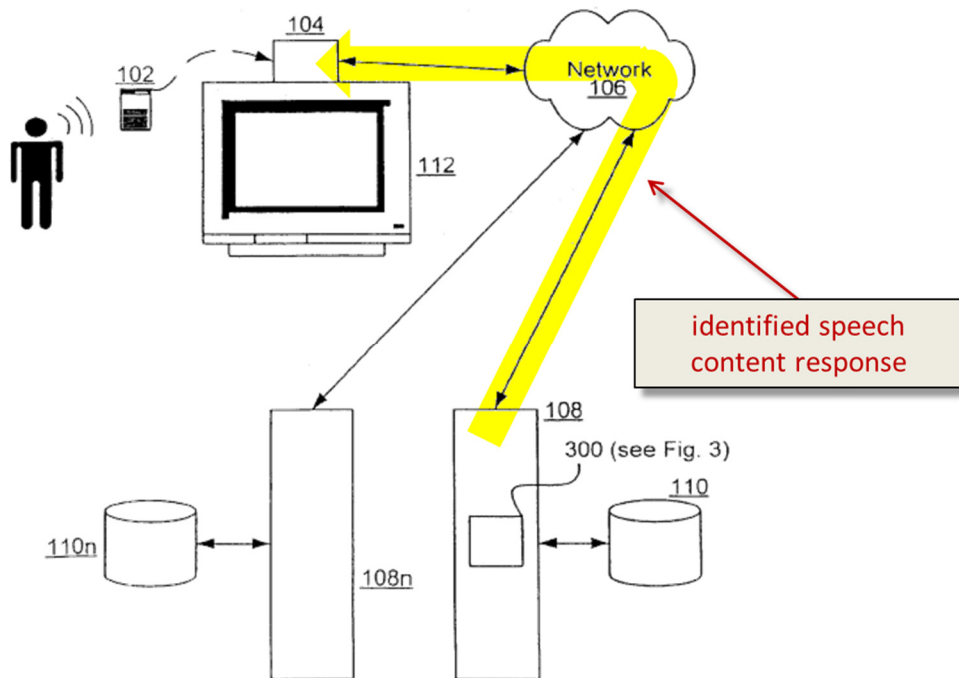
328. Thus, the remote server 108 performs “processing said multiplicity of said received identified speech channels” (i.e., voice commands from multiple different users) by using the processing logic 300 at remote server 108 to “create a multiplicity of identified speech content” (i.e., the identity of the requested content), as recited in claim 1.

329. The final step of claim 1 is:

responding to said identified speech content to create an identified speech content response that is unique, for each of said multiplicity of identified speech contents.

330. Julia discloses this element of claim 1. In Julia, after speech

processing is performed to recognize the spoken request for particular content (i.e., said identified speech content), the remote server retrieves the desired content and electronically transmits it to the user via communications network 106. Julia at 2:27-41, 2:61-67, 4:13-17, 11:60-67 (Ex. 1012).



Julia at Fig. 1a (annotated). In other words, the remote server “responds” to the identified speech content by retrieving the requested programming content and transmitting it to the particular viewer who requested the particular content. Thus, Julia responds to the voice data requesting content from each user by creating “an identified speech content response” (requested programming content) that is “unique” in that it responds to the specific spoken request of a particular viewer. Julia thus discloses this element of claim 1.

2. Claim 2

331. Claim 2 is a dependent claim that adds several limitations. Claim 2 begins as follows:

The method of claim 1, wherein at least one of said identified speech channels has an associated user site; and

332. Julia discloses this limitation. As discussed in connection with claim 1, the “identified speech channels” are the voice requests issued from the individual users, which are captured by a voice input device, such as a microphone, transmitted from the input device to a communications box and then transmitted from the communications box to the remote server. See Paragraphs 317, 319; Julia at 3:39-55 (Ex. 1012). The communications box is “local to the user’s environment,” “coupled to communications network 106,” and can be a “set-top box or a similar communications device that is capable of retransmitting the raw voice data and/or processing the voice data.” Julia at 3:49-54. In other words, the communications box is located at a particular user site. Indeed, this is how the ’196 Patent describes user sites: “Each user site contains a Set Top Box, such as STB 180,” ’196 Patent at 2:8; see also 4:18-19 (“As in FIG. 1, each user site contains a Set Top Box, such as STB 180”) (Ex. 1001). Because each identified speech channel comes from a communications box, which is located at a user site, “at least one of said identified speech channels has an associated user

site,” as recited in claim 2.

333. Claim 2 then adds the following additional limitation:

wherein the step partitioning said received back channel is further comprised of the step of: partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site;

334. Julia discloses this limitation. This part of claim 2 is similar to the “partitioning” step of claim 1 but includes the additional limitation that the received identified speech channels be “from said associated user site.” In Julia, voice requests are transmitted from individual communications boxes to the remote server via network 106 (i.e., the “received back channel”). Julia at 3:39-55, 6:12-26 (Ex. 1012). As discussed in connection with the “partitioning” step of claim 1, the combinations of Julia with Nazarathy or Quigley disclose multiplexing the voice requests for transmission to the remote server via the backchannel, and then demultiplexing (i.e., partitioning) the signals at the server to identify the requests from individual users. See Paragraphs 314-320, above. Thus, the “received back channel” (i.e., upstream channel of network 106) is partitioned “into a multiplicity of said received identified speech channels from said associated user site” (i.e., communications box 104), as recited in the claim.

335. Each “received identified speech channel” from a particular “associated user site” may be further partitioned as well. For instance, a viewer

may speak multiple, separate voice commands. For example, the user could say, “I want to see that movie starring and directed by Clint Eastwood. Can’t remember the title.” Julia at 11:31-32 (Ex. 1012). The user could then review the list of movies where Clint Eastwood is both the star and director. Julia at 11:46-49. The user can then say, “Let’s see Unforgiven.” In this case, the signal from the user site would be partitioned based on the two separate spoken commands.

336. Further, a person of ordinary skill would also recognize that even the individual signals from each local processing unit are commonly divided into discrete portions for transmission. Quigley expressly discloses that, in one embodiment, the cable modem “fragments the data to be transmitted into a plurality of segments if the requested amount is larger than the granted [bandwidth] amount.” Quigley at 1:65-2:1 (Ex. 1014). In Quigley, the segments are then transmitted to the headend using the disclosed TDMA technique. Quigley at 2:2-29, 9:56-10:3; see also 46:32-40 (“The upstream channel 491, is divided into a plurality of time intervals 110, each of which may optionally be further subdivided into a plurality of sub-intervals 489. The upstream channel 491 is thus partitioned so as to facilitate the definition of time slots, such that each of a plurality of cable modems 12 may transmit data packets to the cable modem termination system 10 without interfering with one another, e.g., without having data collisions due to data packets being transmitted at the same time.”). The

headend then reassembles the segments to create each particular signal “from said associated user site.” In order to reassemble the segments to create the particular signal from said associated user site, the headend must first divide the received segments into multiple portions that are each identified specific user sites, and then reassemble those portions that are identified with the same user site together. In this process, the headend “divides” the “received back channel” (the received segments) into “multiple portions that are each identified” with data “originating from system users,” and thus satisfies the construction of “partitioning” as that term is used in the ’196 Patent. See Section Paragraph 87.

337. Thus, Julia in combination with either Nazarathy or Quigley discloses “partitioning said received back channel into a multiplicity of said received identified speech channels from said associated user site,” as recited in claim 2.

338. Claim 2 then adds the following additional limitation:

wherein the step processing said multiplicity of said received identified speech channels is further comprised the step of: processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech contents; and

339. Julia discloses this limitation. After receiving voice requests from individual users, the remote server processes the individual voice requests by performing speech recognition on each to identify what content each user is

requesting. See Paragraphs 322-328, Julia at 3:61-65, 7:5-9:55 (Ex. 1012). Thus, the remote server “processes” the “multiplicity of said received identified speech channels from said associated user site” (i.e., vocal requests transmitted by communications boxes) by performing speech recognition on each of these channels “to create a multiplicity of said identified speech contents” (i.e., the identity of the requested programming).

340. Claim 2 then adds the following additional limitation:

**wherein the step responding to said identified speech content is further comprised the step of:
responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.**

341. Julia discloses this limitation. After performing speech recognition on the user’s voice request (i.e., the identified speech content), the remote server 130 retrieves the requested content and transmits it to the user for viewing on client display device 112 (i.e., the user site). Julia at 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67 (Ex. 1012). Thus, the remote server “responds” to the identified speech content for a specific user site (the location of the communications box 104) to “create said identified speech content response” (the requested content) and transmits it back to the “associated user site,” as recited in the claim.

3. Claim 4

342. Claim 4 is a dependent claim that adds the following limitation:

**The method of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
processing said identified speech content response to create said identified user site response; and sending said identified user site response to said identified user site.**

343. Julia discloses this added limitation. In particular, Julia discloses that “[a]t remote server 108, the voice data is processed by request processing logic 300 in order to understand the user’s request and construct an appropriate query or request for navigation of remote data source 110.” Julia at 3:61-64 (Ex. 1012). “Once the desired information has been retrieved from data source 110, it is electronically transmitted via network 106 to the user for viewing on client display device 112.” Julia at 4:18-20. Thus, Julia discloses processing the “identified speech content response to create said identified user site response” by retrieving the desired content from data source 110 and then “sending said identified user site response to said identified user site” (i.e., transmitting the requested content to the communication box 104, which is coupled to client display device 112), as recited in the claim. See also Julia at 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67.

4. Claim 5

344. Claim 5 is a dependent claim that adds the following limitation:

**The method of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and**

345. Julia discloses this limitation. Julia notes that users can request content, which can include movies-on-demand, and teaches a natural language voice input system to take “full advantage” of that content. Julia at 1:30-43 (Ex. 1012). Julia teaches providing multimedia content in response to those requests, which, in the case of movies-on-demand, would create a financial consequence for the user (i.e., a purchase by the user). See Julia at 3:61-4:17. Indeed, Julia states that the disclosed invention could be implemented “on any number of different hardware and software computing platforms and environments” specifically including the “Diva Systems video-on-demand system.” Julia at 6:47-59. A person of ordinary skill in the art would have known that the Diva System provided pay-per-view functionality. Thus, Julia discloses to one of ordinary skill in the art requesting pay-per-view or video-on-demand content by issuing voice commands, which the system would process, and providing responsive content. A person of ordinary skill would also understand that pay-per-view and video-on-demand functionality, such as that provided by the Diva system, would require a

purchase, resulting in a bill. Thus, Julia discloses “assessing said speech response” (i.e., assessing the content requested by the user) associated with a particular “user site” (i.e., the location of the communications box 104) to “create a financial consequence” (i.e., payment for a video-on-demand movie), as recited in the claim.

346. The last step of claim 5 requires:

billing said user site based upon said financial consequence.

347. Julia could be combined with the teaching of numerous other prior art references that disclose pay-per-view movie functionality in cable and other television networks. By the end of the 1990s, pay-per-view systems had been implemented by a number of different cable operators and the concept was well known. A person of ordinary skill in the art would have been familiar with the basic functionality of a pay-per-view system, which generally follows the steps recited in claim 5 (and also claims 6, 18, 19, 31, 32, and 65) of the '196 Patent. Specifically, a typical prior art pay-per-view movie system would present the user with a list of available titles and either the associated pricing or at least a warning that selecting a title from the list would lead to a charge—hence the name “pay-per-view.” When the user selected a movie to watch, the system would commonly ask the user to confirm the purchase (for example by pressing a particular button on the television remote control). Once the user confirmed the purchase, the pay-per-view system would deliver the requested movie for viewing and bill the user,

commonly by adding the purchase price to the subscriber's monthly bill.

348. A person of ordinary skill in the art would have understood these basic features of a pay-per-view system, like the one disclosed in Julia, because they were well known in the prior art. For example, U.S. Patent No. 5,477,262 is a prior art patent entitled "Method and Apparatus for Providing an On-Screen User Interface for a Subscription Television Terminal" that describes the basic steps of permitting a user to purchase pay-per-view content. Banker at 16:62-17:3, Fig. 6F (Ex. 1015). In the Banker system, a user can purchase a pay-per-view movie by pressing the "BUY key" on the television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 ("When the terminal of FIG. 3 is in the On mode and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by actuating the BUY key initiates a buy sequence. Consequently, this key is used to purchase an event."). Once entered by the user, the "code is checked with the code stored in memory" (i.e., "assessing a response identified" with the user, as recited in claim 8). "If correctly entered, the subscriber is asked to press the BUY key again ... and the subscriber is thanked" Banker at 17:1-3 . The user is then billed for the purchase. Banker at 7:60-63 ("Billing computer 101 includes a subscriber database and generates a monthly bill for the subscribers in the system based on level of service and any

pay-per-view and impulse pay-per-view purchases.”). Thus, Banker discloses the “assessing” and “billing” steps of claim 5. Moreover, as explained below in paragraphs 446-447, a person of ordinary skill in the art would have been motivated to combine Banker with Julia. Accordingly, Julia alone or in combination with Banker discloses the elements of claim 5 and renders the claim obvious.

349. As another example, U.S. Patent No. 6,314,573 is a prior art patent entitled “Method And Apparatus For Providing Subscription-On-Demand Services For An Interactive Information Distribution System” that also describes the basic steps of permitting a user to purchase pay-per-view content using a different platform. Gordon at 8:66-9:63, Fig. 3B (Ex. 1016). The Gordon patent was assigned to Diva Systems, which offered a subscription pay-per-view platform in the 1990s. The Gordon system offers the subscriber an option to purchase an on-demand program subscription. Gordon at 9:64-10:9, Figs. 3B and 8. If the user elects to make a purchase, the system requests a master PIN as a confirmation. Gordon at 10:16-27. If the PIN number is valid, the system updates the billing system with the new subscriber’s account number. Gordon at 10:43-46. (“After updating the subscription data base, the routine updates, at step 382, the administrative system such that the interactive information distribution system's billing system is updated with the new subscriber's account number.”); Gordon at

2:60-63 (“As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that package and is billed accordingly.”). Once a user subscribes to a package of programming and pays the fixed price, he can view any program within the package at any time. Gordon at 3:61-63. As explained below, persons of ordinary skill in the art would have been motivated to combine Julia with Gordon. See Paragraphs 446-447. Accordingly, Julia alone or in combination with Gordon discloses the elements of claim 5 and renders the claim obvious.

5. Claim 6

350. Claim 5 is a dependent claim that adds the following limitation:

**The method of claim 2, wherein the step responding to said identified speech contents from said associated user site is comprised of the steps of:
assessing said speech response to create a financial consequence identified with said user site;**

351. Julia discloses this limitation. As discussed above in connection with claim 8, Julia notes that users can request content, which can include movies-on-demand, and teaches a natural language voice input system to take “full advantage” of that content. Julia at 1:30-43 (Ex. 1012). Julia teaches providing multimedia content in response to those requests, which, in the case of movies-on-demand, would create a financial consequence for the user (i.e., a purchase by the user).

See Julia at 3:61-4:17. Indeed, Julia states that the disclosed invention could be implemented “on any number of different hardware and software computing platforms and environments” specifically including the “Diva Systems video-on-demand system.” Julia at 6:47-59. A person of ordinary skill in the art would have known that the Diva System provided pay-per-view functionality. Thus, Julia discloses to one of ordinary skill in the art requesting pay-per-view or video-on-demand content by issuing voice commands, which the system would process, and providing responsive content. A person of ordinary skill would also understand that pay-per-view and video-on-demand functionality, such as that provided by the Diva system, would require a purchase, resulting in a bill. Thus, Julia discloses “assessing” a response (i.e., assessing the content requested by the user) associated with a particular “user device” (i.e., voice input device or communications box) to “create a financial consequence” (i.e., payment for a video-on-demand movie), as recited in the claim.

352. In addition, Banker also discloses this limitation of the claim. In Banker, a user can purchase a pay-per-view movie by pressing the “BUY key” on the television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (“When the terminal of FIG. 3 is in the On mode and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by

actuating the BUY key initiates a buy sequence. Consequently, this key is used to purchase an event.”) (Ex. 1015). Thus, Banker discloses “assessing” a response (i.e., assessing the user’s key entries) to “create a financial consequence” (i.e., purchase of a pay-per-view movie), as recited in the claim. The combination of Julia with Banker also discloses this element.

353. Gordon also discloses this limitation of the claim. Gordon’s Figure 3B illustrates a flow chart of the basic steps in ordering video content. The user selects a category from a category menu, and then a title from a title menu (e.g. Sesame Street). Gordon at 9:18-28; Fig. 3B (Ex. 1016). After receiving a title selection, the system determines the title price. Gordon at 9:29-31; Fig. 3B. Hence, the system assesses the user’s response (i.e. choice of title) to create a financial consequence, i.e. the title price. Thus, Gordon discloses “assessing” a response (i.e., assessing the user’s choice of a title) to “create a financial consequence” (i.e., determine the price of the title), as recited in the claim.

354. The next step of claim 6 requires:

displaying said financial consequence to create a displayed financial consequence at said user site;

355. As discussed in connection with claim 5, Julia explains that a viewer can use voice commands to select and order pay-per-view movies. Julia at 1:30-43, 10:12-13, 11:31-32. (Ex. 1012). A person of ordinary skill in the art would

know that a pay-per-view movie purchase involves informing the user that they are making a purchase, such as by displaying the price of the requested programming to the user. For example, the Banker patent describes a prior art pay-per-view system that presented the user with a screen stating: “NEXT, PLEASE PRESS [BUY] TO PURCHASE THIS EVENT” to confirm the viewer’s intent to accept the purchase (i.e., the “financial consequence”). Banker at 16:62-17:3, Fig. 6F (Ex. 1015). As another example, the Gordon patent illustrates at Fig. 3B a flow chart of the basic steps in ordering video content. Once the user selects a title and the price has been determined, the system queries the user to see if he wants to pay “a la carte” to view his selection. 9:56-63 (“...the video session manager sends, at step 352, an applet for a second type of title information screen. At step 354, the terminal creates a screen informing the subscriber of the reason for the non-zero price and also presents the user with an “a la carte” purchase option. The subscriber will then be able to purchase and view the a la carte selection or return to a previous menu.”) Thus, either Julia alone or combined with Banker or Gordon discloses this limitation of claim 6.

356. Next, claim 6 requires:

confirming said displayed financial consequence from said user site to create a financial commitment; and

357. As discussed in connection with claim 5, Julia explains that a viewer

can use voice commands to select and order pay-per-view movies. Julia at 1:30-43, 10:12-13, 11:31-32 (Ex. 1012). A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase and confirming their intent to do so. For example, the Banker patent describes a prior art pay-per-view system that requires the user to press the BUY key, enter a system provided code, and then asking the subscriber to “press the BUY key again” to confirm the intent to purchase (i.e., the “financial consequence”) to “create a financial commitment,” as claimed. Banker at 16:62-17:3, Fig. 6F (Ex. 1015). As noted above, the Gordon patent allows a viewer to pay “a la carte” to view a selection. Gordon at 9:56-63 (“...the video session manager sends, at step 352, an applet for a second type of title information screen. At step 354, the terminal creates a screen informing the subscriber of the reason for the non-zero price and also presents the user with an “a la carte” purchase option. The subscriber will then be able to purchase and view the a la carte selection or return to a previous menu.”) (Ex. 1016). Gordon provides an example in Figure 8, which gives the viewer the option to purchase a program subscription by clicking on-screen buttons. Gordon at Fig. 8; see also 6:43-47 (on-screen buttons are selected via the remote control). Thus, either Julia alone or combined with Banker or Gordon discloses this limitation of claim 6.

358. The following figure is from Banker:

SCREEN 7a
IPPV BUY

NEXT, PLEASE PRESS
[BUY]
TO PURCHASE THIS EVENT

OR

PRESS [MENU] TO EXIT.

Banker at Fig. 6F (Ex. 1015) (excerpted).

359. The following figure is from Gordon:

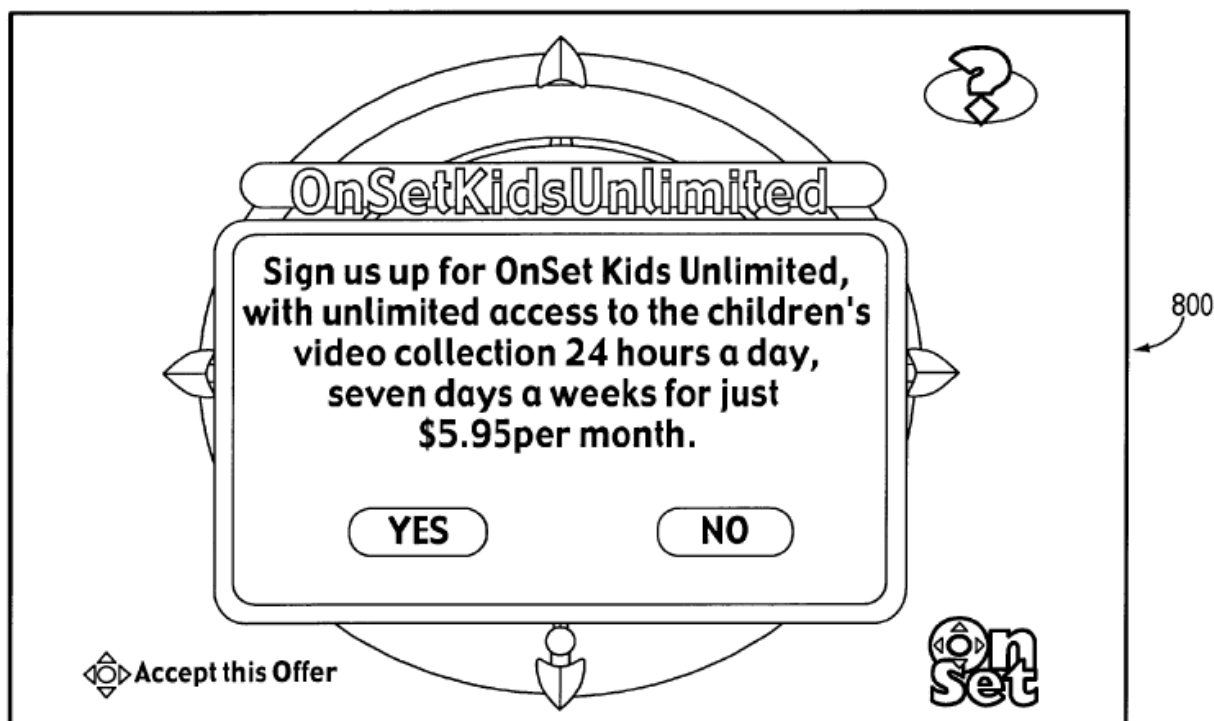


FIG. 8

Gordon at Fig. 8 (Ex. 1016).

360. Claim 6's last requirement is:

billing said user site based upon said financial commitment.

361. Julia discloses this limitation. As discussed above, Julia discloses pay-per-view and video-on-demand functionality. Julia at 1:30-43, 10:12-13, 11:31-32 (Ex. 1012). Also as discussed above, a person of ordinary skill would also understand that pay-per-view and video-on-demand functionality, such as that provided by the Diva system, would require a purchase, resulting in a bill. Banker also discloses billing the subscriber for a pay-per-view purchase. Banker at 7:60-63, 16:62-17:3 ("generate a monthly bill ... based on level of service and any pay-per-view and impulse pay-per-view purchases") (Ex. 1015). Gordon also discloses billing the subscriber for a pay-per-view purchase. Gordon at 10:39-47 ("If the PIN is valid [] the routine updates, at step 382, the administrative system such that the interactive information distribution system's billing system is updated with the new subscriber's account number."); 2:60-63 ("As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that package and is billed accordingly.") (Ex. 1016). Thus, Julia alone or combined with Banker or Gordon discloses "billing said user site based upon said financial commitment," as claimed.

6. Claim 12

362. Claim 12 is a dependent claim that adds the following limitation:

The method of claim 2, wherein the step responding to said identified speech content from said associated user site is further comprised of the step of:
responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech content identified with said user site.

363. Julia discloses this limitation. The claim recites responding to the speech content identified as to said user site (i.e., voice requests) “based upon natural language.” ’196 Patent at 52:53 (Ex. 1001). Julia discloses performing speech recognition “based upon natural language,” as recited in the claim. Indeed, the entire disclosure of Julia is directed at allowing users to request programming content using natural language requests. See Julia at 1:38-42 (“Allowing *spoken natural language requests* as the input modality for rapidly searching and accessing desired content *is an important objective* for a successful consumer entertainment product”) (emphasis added); 1:61-63 (“In order to deliver a true ‘win’ for users, the voice- driven front-end *must accept spoken natural language input* in a manner that is intuitive to access.”) (Ex. 1012) (emphasis added).

364. In Julia’s “data navigation system driven by spoken natural language input,” (Julia at 3:37-38, Ex. 1012), voice data is first “interpreted in order to understand the user’s request for information” by using a speech recognition engine, and then is linguistically parsed and interpreted using a “natural language

interpreter (or parser) 320.” Julia at 7:5-7, 7:47-53. Julia provides a detailed description of how voice requests are processed to identify content specified by a user, and includes examples of natural language requests, such as “Show me the movie ‘Unforgiven’” or “I want to see that movie starring and directed by Clint Eastwood. Can’t remember the title.” See Julia at 10:12-13, 11:31-32; see generally Julia at 6:61-12:62.

365. After performing speech recognition on the user’s voice request (i.e., said speech content identified as to said user site), the remote server 130 retrieves the requested content and transmits it to the user for viewing on client display device 112 (i.e., the user site). Julia at 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67 (Ex. 1012). The remote server thus “responds” to the speech content identified for each user site (i.e., the location of the communications box 104) by creating a response (i.e., retrieving and transmitting the requested content) to the specific user site, as recited in the claim.

366. Thus, Julia discloses “responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech content identified with said user site,” as recited in claim 12.

7. Claim 13

367. Claim 13 is a dependent claim that adds the following limitation:

The method of claim 2, wherein the step processing said multiplicity of said received speech channels is further comprised

**for at least one of said user sites is further comprised of the step of:
processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.**

368. Julia discloses this added limitation. As discussed in connection with claim 12, Julia discloses speech recognition processing “based upon natural language.” See Paragraph 363-366, above. In particular, Julia’s entire disclosure is directed to accessing programming content using natural language spoken requests. Julia at 1:38-42, 1:61-63 (Ex. 1012). In Julia’s system, voice data is first interpreted by a speech recognition engine, and then linguistically parsed and interpreted using a natural language interpreter. Julia at 7:5-7, 7:47-53. Julia provides a detailed description of how voice requests are processed to identify content specified by a user. See generally Julia at 6:61-12:62.

369. Further, Julia discloses processing the voice requests “to create said speech content identified as to said user site,” as recited in the claim. Julia discloses that the remote server 108 performs speech recognition to identify the requested content. Julia at 2:27-41, 2:61-67, 4:10-30, 10:12-13, 11:47-67 (Ex. 1012). The identity of the requested content constitutes the claimed “speech content identified as to said user site.”

8. Claim 27

370. The preamble of claim 27 reads as follows:

A system supporting speech recognition in a network, said system comprising:

371. Julia discloses a system supporting speech recognition in a network as claimed in claim 27. Julia discloses a system “for navigating network-based electronic data sources in response to spoken input requests.” Julia at 2:27-30 (Ex. 1012). “When a spoken input request is received from a user, it is interpreted, such as by using a speech recognition engine to extract speech data from acoustic voice signals, and using a language parser to linguistically parse the speech data.” Julia at 2:30-34. The processing of a user’s spoken input request is “processed by request processing logic 300,” which is stored in “remote server 108.” Julia at 3:61-64. The remote server is connected to users’ set top boxes (called “communications box 104” in Julia) via network 106, which “may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.” Julia at 3:49-55, 4:31-35; Fig. 1a. The remote server is also capable of handling simultaneous voice requests from multiple users. Julia at 6:12-26. After performing speech recognition to understand a user’s request, the remote server retrieves the desired content and transmits it to the viewer. Julia at

4:18-20, 11:60-67. The remote server 108 containing processing logic 300 is therefore a “system supporting speech recognition in a network” as the preamble of claim 27 requires.

372. The first element of claim 27 reads as follows:

a speech recognition system coupled to a wireline node in said network for receiving a back channel from a multiplicity of user sites coupled to said network, further comprising

373. Julia discloses this element. As discussed above, the “speech recognition system” is the remote server 108, which includes processing logic 300. The remote server 108 is also a “wireline node” as that term is used in the ’196 Patent because it is a “a network node providing video or cable television delivery to multiple users over physical wires between the node and the multiple user locations.” See Paragraphs 80-81 above. For example, the network server 108 delivers videos to users over network 106. Julia at 11:60-65 (example of delivering the movie Unforgiven to a user) (Ex. 1012). As shown in Fig. 1a below, network 106 connects the network server 108 with the client display devices 112, and may be embodied in cable television lines, fiber-optic cable, or other types of “hardwired” (i.e., physical) connections. Julia at 4:31-35.

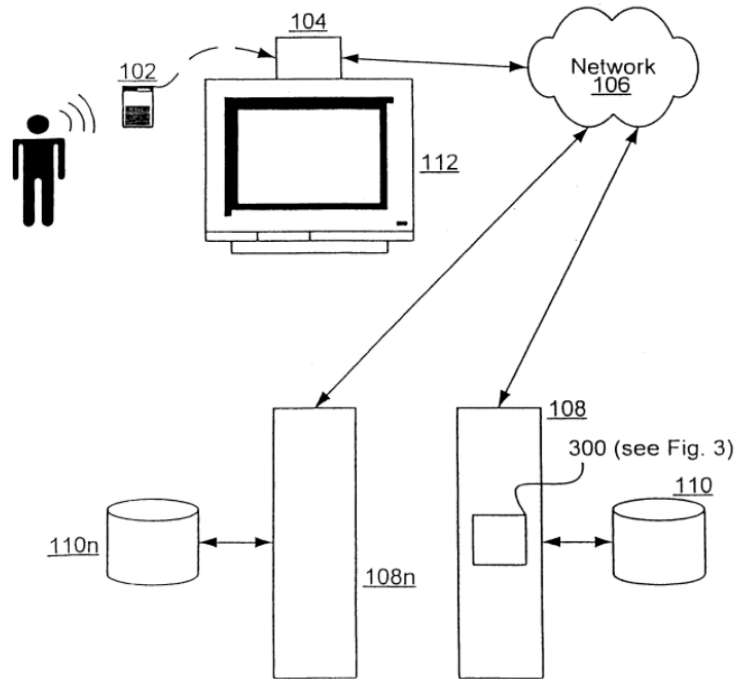
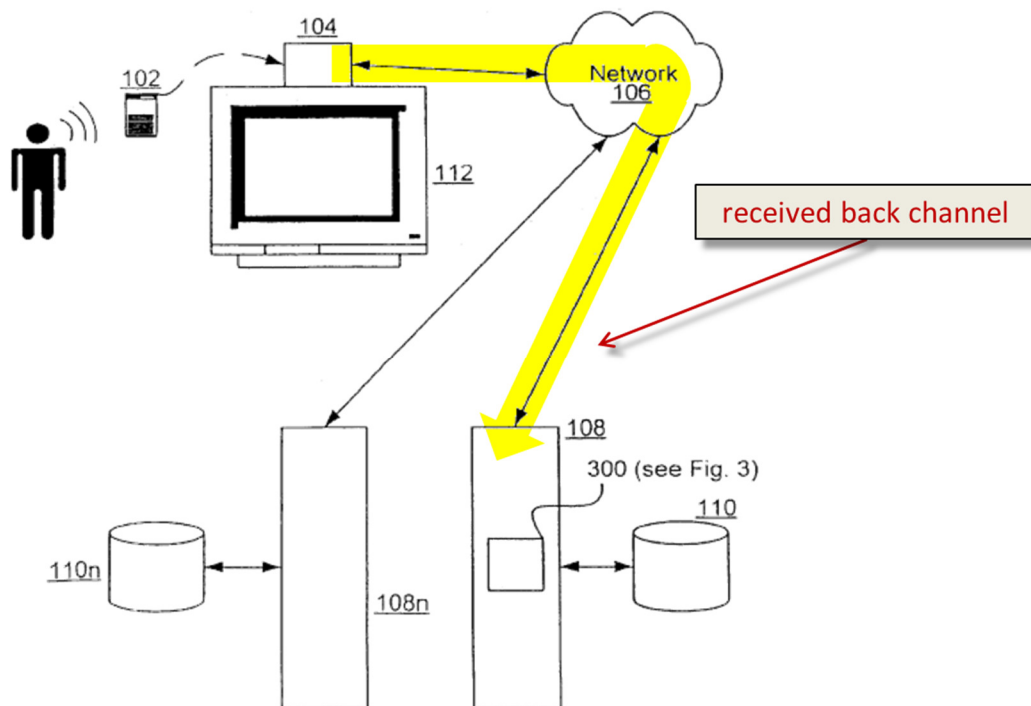


Fig. 1a

Julia at Fig. 1a. The remote server 108 is also connected with multiple users. Julia at 6:12-26. The network server 108 is thus a “wireline node,” as that term is used in the ’196 Patent, that is “in a network supporting at least one of cable television delivery and video delivery,” as recited in the claim.

374. Julia also discloses that the remote server 108 “receiv[es] a back channel from a multiplicity of user sites coupled to said network” as this claim requires. A back channel is an “upstream communication channel delivering signals from multiple user sites to a central wireline node.” See Paragraph 82-83. Julia discloses that a user’s voice input data is transmitted from the user’s “communication box 104 (e.g., a set-top box or a similar communications device

that is capable of retransmitting the raw voice data and/or processing the voice data)” through network 106 to remote server 108, where the voice data is processed, as shown in Fig. 1a. Julia at 3:48-64, Fig. 1 (Ex. 1012).



Julia Fig. 1a (annotated). “Network 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.” Julia at 4:31-35. Julia also discloses that multiple network users may issue requests “simultaneously or otherwise,” and that the data source 110 and processing logic are “preferably designed and implemented to support queuing and multi-tasking of requests from multiple simultaneous network users, as will be appreciated by those of skill in the

art.” Julia at 6:12-26. In other words, multiple users may send voice input data simultaneously through network 106 to remote server 108. The remote server 108 (the “speech recognition system coupled to a wireline node”) therefore “receiv[es] a back channel” as recited in the claim.

375. Julia also discloses that the back channel is from a multiplicity of user sites coupled to the network. As discussed above, the remote server receives voice requests from individual communications boxes 104. The communications box is “local to the user’s environment,” “coupled to communications network 106,” and can be a “set-top box or a similar communications device that is capable of retransmitting the raw voice data and/or processing the voice data.” Julia at 3:49-54 (Ex. 1012). In other words, the communications box is located at a particular user site. Indeed, this is how the ’196 Patent describes user sites: “Each user site contains a Set Top Box, such as STB 180,” ’196 Patent at 2:8; see also 4:18-19 (“As in FIG. 1, each user site contains a Set Top Box, such as STB 180”) (Ex. 1001). Because users’ speech requests each come from a communications box, which is located at a user site, the back channel is from “a multiplicity of user sites coupled to the network” as recited in claim 27.

376. The second element of claim 27 reads as follows:

a back channel receiver, for receiving said back channel to create a received back channel;

377. This element is similar to the first element of claim 1, with the exception that while claim 1 simply requires “receiving said back channel to create a received back channel,” this element requires “*a back channel receiver*, for receiving said back channel to create a received back channel.” Compare ’196 Patent at 51:1-2 with ’196 Patent at 55:15-16 (claim 27) (Ex. 1001) (emphasis added). As discussed in connection with claim 1 and in connection with the previous limitation, Julia discloses that the remote server 108 receives a back channel to create a received back channel. See Paragraphs 307-312, 374-375; Julia at 3:48-64, 4:31-35, 6:12-26, Fig. 1 (Ex. 1012). The remote server 108 is thus the “back channel receiver” recited in claim 27.

378. The third element of claim 27 reads as follows:

a speech channel partitioner, for partitioning said received back channel into a multiplicity of received identified speech channels; and

379. Julia in combination with Quigley or Nazarathy discloses this element. This element is similar to the second element of claim 1, with the exception that while claim 1 simply requires “partitioning said received back channel into a multiplicity of received identified speech channels,” this element requires “*a speech channel partitioner*, for partitioning said received back channel into a multiplicity of received identified speech channels.” Compare ’196 Patent at 51:1-2 with ’196 Patent at 55:15-16 (claim 27) (Ex. 1001) (emphasis added). Julia

discloses that multiple network users may issue voice request simultaneously, and the system is “preferably designed and implemented to support queuing and multi-tasking of requests from multiple simultaneous network users, as will be appreciated by those of skill in the art.” Julia at 6:12-26 (Ex. 1012). Julia does not, however, provide any details regarding how the system can be designed to support such queuing and multi-tasking of requests from multiple simultaneous network users.

380. As discussed in detail with respect to claim 1, persons of ordinary skill in the art knew that one way to support the simultaneous transmission and processing of voice requests from multiple network users is to multiplex the voice requests together for transmission over the back channel, and then demultiplex (partition) the voice requests at the remote server. Such multiplexing and demultiplexing is disclosed by both Nazarathy and Quigley. For the reasons discussed in connection with the second element of claim 1, the combinations of Julia with Nazarathy and Quigley thus disclose this element of claim 27, where the remote server 108 is the “speech channel partitioner.” See Paragraphs 314-320.

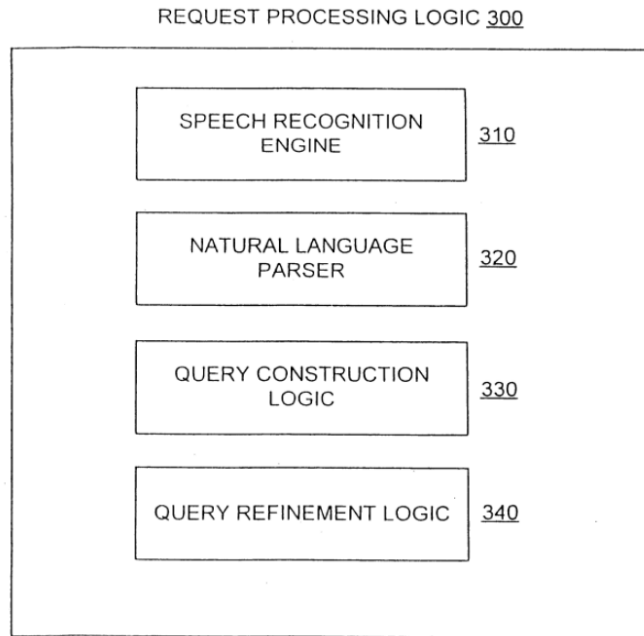
381. The fourth element of claim 27 reads as follows:

a processor network executing a program system comprised of program steps residing in memory accessibly coupled to at least one computer in said processor network;

382. Julia discloses this element. The “program steps” this element refers

to are (1) “processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content;” and (2) “responding to said identified speech content to create an identified speech content response, for each of said multiplicity of said identified speech contents,” which follow immediately after this limitation. ’196 Patent at 55:24-31 (Ex. 1001). As explained below in connection with those elements, the remote server 108 processes the multiplicity of identified speech channels (the user’s voice commands) by performing speech recognition on the voice commands to identify the programming content requested by the user. See Paragraphs 388, below. Once identified, the remote server then responds by transmitting the requested programming content to the user. See Paragraphs 390, below.

383. The remote server includes “processing logic 300,” which performs speech recognition on the user’s commands. Julia at 3:61-62 (Ex. 1012). Processing logic 300 includes a number of components, as shown in Figure 3 of Julia:



Julia Fig. 3

384. A person of ordinary skill in the art understands that processing logic 300, in addition to all of its components, must necessarily “reside” on memory “accessibly coupled” to the remote server 108. Otherwise, remote server 108 would not be able to execute these engines and logic. Julia further discloses that remote server 108 may consist of multiple servers. Julia at 3:54-55 (Ex. 1012). When implemented using multiple servers, a “processor network” executes the processing and responding steps, which reside in memory accessibly coupled to at least one computer, as recited in this claim. A person of ordinary skill in the art understands that a “computer,” such as a remote server computer, includes a processor that executes steps residing in memory accessibly coupled to that computer. Thus, if multiple computers are used for the speech processing step,

then at least two processors, which are coupled together, are involved. The processors of these two computers form a “processing network” as recited in the claim.

385. Julia also discloses that data source 110, which stores video content, “preferably resides on a central server or servers—which may or may not be the same as server 108.” Julia at 4:4-13 (Ex. 1012). Julia thus discloses that data source 110 may consist of multiple servers. The remote server (or servers) 108 retrieves the desired content from data source 110, and then electronically transmits it via network 106 to the user. Data source 110, which stores its content on memory, is also therefore accessibly coupled to remote server 108, and is part of a “processing network” along with remote server 108 that together executes the processing and responding steps discussed below.

386. In sum, the “processor network” executing the processing and responding steps of claim 27 can consist of (1) multiple remote servers 108; (2) multiple data source 110 servers; and (3) a combination of remote servers 108 and data source 110 servers. Each server contains “program steps residing in memory” accessibly coupled to at least one computer (namely, the servers) in the processor network, as required by this claim.

387. The fifth element of claim 27 reads as follows:

wherein said program system is comprised of the program steps

of:

processing said multiplicity of said received identified speech channels to create a multiplicity of identified speech content;

388. This element is identical to the third element of claim 1. See Paragraph 59 above. Julia discloses this limitation for the reasons discussed in connection with claim 1. See Paragraphs 322-328, above. In sum, after receiving the voice input data from various users, the remote server 108 processes the voice input data using processing logic 300. Julia at 3:61-66 (Ex. 1012). Processing logic 300 performs speech recognition on the voice input data to extra words from the voice data, and then performs natural language parsing of those words to understand the users' voice data. See Paragraphs 322-328, above; see generally Julia at 7:1-10:30. Based on that understanding, processing logic 300 identifies the content requested by the users—the multiplicity of identified speech content. Julia at 2:27-41, 2:61-67, 4:10-30, 10:12-13, 11:47-67. Thus, the remote server 108 performs “processing a multiplicity of said received identified speech channels” (i.e., voice requests from multiple different users) by using the processing logic 300 at remote server 108 to “create a multiplicity of identified speech content” (i.e., the identity of the requested content), as recited in claim 27.

389. The sixth element of claim 27 reads as follows:

responding to said identified speech content to create an identified

speech content response, for each of said multiplicity of said identified speech contents; and

390. This element is nearly identical to the fourth element of claim 1. See Paragraph 59 above. Julia discloses this limitation for the reasons discussed in connection with claim 1. See Paragraph 330, above. In short, after identifying the content requested by each user, the remote server 108 retrieves and transmits the requested content to the requesting user. Julia at 2:27-41, 2:61-67, 4:13-17, 11:60-67 (Ex. 1012).

391. The last element of claim 27 reads as follows:

wherein said network supports at least one of the collection comprising: cable television delivery to said multiplicity of user sites; and video delivery to said multiplicity of user sites.

392. Julia discloses this element. Julia discloses that network 106 can be used to supply cable television and video-on-demand. Julia at 4:31-35, 10:9-18, 10:44-55, 11:16-45 (Ex. 1012). Julia further discloses that the network can be used to deliver content to a multiplicity of users. Julia at 6:12-26. Julia thus discloses that its network supports both “cable television delivery to said multiplicity of user sites” and “video delivery to said multiplicity of user sites” as recited by the claim.

9. Claim 28

393. Claim 28 is a dependent claim that adds the following limitations:

The system of claim 27, wherein at least one of said identified

speech channels has an associated user site; and

394. Other than minor differences in their preambles (claim 2 recites a “method” whereas claim 28 recites a “system), this element is identical to the first element of claim 2. See Paragraph 60, above. Julia discloses this limitation for the reasons discussed in connection with claim 2. See Paragraph 332, above. In short, the ’196 Patent states that a “user site” contains a set top box. ’196 Patent at 2:8, 4:18-19. (Ex. 1001). Julia discloses that the “identified speech channels”—*i.e.*, users’ voice requests—each come from various “network users” that may issue requests “simultaneously or otherwise.” Julia at 6:12-26 (Ex. 1012). Users speak their commands into a voice input device 102, which is coupled to a “communications box 104,” such as a set-top box or similar device, that is “local to the user’s environment” (*i.e.*, at a user site). Julia at 3:39-54.

395. The next element of claim 28 recites:

**wherein partitioning said received back channel is further
comprised of:**

**partitioning said received back channel into a multiplicity of said
received identified speech channels from said associated user site;**

396. This limitation is substantively the same as the corresponding limitation recited in claim 2. See Paragraph 60, above. Julia in combination with Nazarathy and Quigley, disclose this limitation for the reasons discussed in connection with claim 2. See Paragraphs 334-337, above. In short, Julia discloses

that its system should be designed to support simultaneous requests from multiple users. Julia at 6:12-26. The combinations of Julia with Nazarathy or Quigley disclose multiplexing the voice requests for transmission to the remote server via the backchannel, and then demultiplexing (i.e., partitioning) the signals at the server to identify the requests from individual users. See Paragraph 334, above. Thus, the “received back channel” (i.e., upstream channel of network 106) is partitioned “into a multiplicity of said received identified speech channels from said associated user site” (i.e., communications box 104), as recited in the claim.

397. In addition, each “received identified speech channel” from a particular “user site” may be further partitioned. For example, a user located at one user site may issue multiple commands. Each command issued is demultiplexed—partitioned—by the remote server computer. See Paragraph 335, above.

398. Finally, each individual signal from each local processing unit may itself be divided into discrete portions for transmission. For example, Quigley discloses that in systems that use TDMA, fragments of data can be transmitted in a plurality of segments. The headend then reassembles the segments to create each particular signal from each user site, as claimed. See Paragraph 336, above. The combination of Julia and Quigley thus also render this element obvious.

399. The next element of claim 28 recites:

wherein the program step processing said multiplicity of said received identified speech channels is further comprised of the program step of:

processing said multiplicity of said received identified speech channels from said associated user site to create a multiplicity of said identified speech content; and

400. This limitation is the same as the corresponding limitation recited in claim 2, with the exception that while claim 2 recites a “step,” this element recites a “program step.” See Paragraph 60, above. Julia discloses this limitation for the reasons discussed in connection with claim 2. See Paragraph 339, above. In short, after receiving voice requests from various users, the remote server 108 of Julia “process[es]” the “multiplicity of said received identified speech channels from said associated user site” (i.e., the voice requests”) by performing speech recognition on them “to create a multiplicity of identified speech content” (i.e., the identity of the requested content). See, e.g., Julia at 3:61-65, 7:5-9:55 (Ex. 1012). This processing is performed every time a user at a user site uses a voice request to request programming content.

401. The next element of claim 28 recites:

wherein the program step responding to said identified speech content is further comprised of the program step of:
responding to said identified speech content from said associated user site to create said identified speech content response for said associated user site.

402. This limitation is the same as the corresponding limitation recited in claim 2, with the exception that while claim 2 recites a “step,” this element recites a “program step.” See Paragraph 60, above. Julia discloses this limitation for the reasons discussed in connection with claim 2. See Paragraph 341, above. In short, the remote server 130 responds to the identified speech content by retrieving the requested content and transmitting it to the user—*i.e.*, the user site. See Julia at 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67 (Ex. 1012).

10. Claim 30

403. Claim 30 is a dependent claim that adds the following limitations:

The system of claim 28, wherein the program step responding to said identified speech contents from said associated user site is comprised of the program steps of:

processing said identified speech content response to create said identified user site response; and sending said identified user site response to said identified user site.

404. Other than minor differences in the preamble (claim 4 recites a “method” whereas claim 30 recites a “system”), this limitation is identical to the corresponding limitation recited in claim 4. See Paragraph 61, above. Julia discloses this limitation for the reasons discussed in connection with claim 4. See Paragraph 343, above. In sum, Julia discloses that the remote server 108 processes voice requests using processing logic 300 to understand the user’s request and construct an appropriate query or request for navigation of the remote data source

110. Julia at 3:61-64 (Ex. 1012). Once the content is identified, it is transmitted from the remote server to the user. Julia at 4:18-20, 11:60-67; see also 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67.

11. Claim 31

405. Claim 31 is a dependent claim that adds the following limitations:

**The system of claim 28, wherein the program step responding to said identified speech contents from said associated user site is comprised of the program steps of:
assessing said speech content response identified as to said user site to create a financial consequence identified as to said user site; and**

406. The “assessing” step of claim 31 is identical to the “assessing” step of claim 5. See Paragraph 62, above. Accordingly, Julia discloses this limitation for the same reasons. See Paragraphs 345, above. In short, Julia teaches providing multimedia content in response to voice requests, which, in the case of movies-on-demand, would create a financial consequence for the user (i.e., a purchase by the user). See Julia at 3:61-4:17 (Ex. 1012). Julia further states that its invention could be implemented on the “Diva Systems video-on-demand system,” which persons of ordinary skill in the art knew provided pay-per-view functionality. Julia at 6:47-59; Paragraph 345, above. Thus, Julia discloses “assessing said speech content response” (i.e., assessing the programming content requested by the user) associated with a particular “user site” (i.e., the location of the communications

box 104) to “create a financial consequence” (i.e., payment for a pay-per-view movie), as recited in the claim.

407. The last element of claim 31 is:

billing said user site based upon said financial consequence.

408. The “billing” step of claim 31 is identical to the “billing” step of claim

5. See Paragraph 62, above. Accordingly, Julia alone or in combination with Banker and Gordon disclose this limitation for the same reasons. See Paragraphs 347-349, above. In short, Julia discloses implementing its system on pay-per-view systems, including the Diva Systems video-on-demand system. Persons of ordinary skill in the art understood that in pay-per-view systems, users are billed for the content they view. See Paragraph 347.

409. In addition, Banker describes a system that allows users to purchase pay-per-view content by pressing the “BUY key” on a television remote and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). After following addition steps, the purchase is confirmed and the user is billed. Banker at 17:1-3, 7:60-63. Similarly, Gordon describes a system that allows a user to purchase an on-demand video subscription. Gordon at 9:64-10:9, Figs. 3B and 8 (Ex. 1016). If the user confirms the purchase, the user is billed for the subscription. Gordon at 2:60-63, 3:61-63. Because persons of ordinary skill in the art would be motivated

to combine Julia with Banker or Gordon (see paragraphs 446-447, below), such combinations would also disclose this limitation.

12. Claim 32

410. Claim 32 is a dependent claim that adds the following limitations:

**The system of claim 28, wherein the program step responding to said identified speech content from said associated user site is comprised of the program steps of:
assessing said speech response to create a financial consequence identified as to said user site;**

411. The “assessing” step of claim 32 is identical to the “assessing” step of claim 6. See Paragraph 63, above. Accordingly, Julia discloses this limitation for the same reasons. See Paragraphs 351-353, above. In short, Julia teaches providing multimedia content in response to voice requests, which, in the case of movies-on-demand, would create a financial consequence for the user (i.e., a purchase by the user). See Julia at 3:61-4:17 (Ex. 1012). Julia further states that its invention could be implemented on the “Diva Systems video-on-demand system,” which persons of ordinary skill in the art knew provided pay-per-view functionality. Julia at 6:47-59, Paragraph 351, above. Thus, Julia discloses “assessing said speech content response” (i.e., assessing the programming content requested by the user) associated with a particular “user site” (i.e., the location of the communications box 104) to “create a financial consequence” (i.e., payment for a pay-per-view movie), as recited in the claim.

412. In addition, both Banker and Gordon also disclose this limitation. See Paragraphs 352-353, above. For example, Banker discloses purchasing a pay-per-view movie by pushing the “BUY key” on a television remote control and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). Banker thus discloses “assessing said speech response” (i.e. assessing the user’s key entries) to “create a financial consequence” (i.e., purchase of a pay-per-view movie), as recited in the claim.

413. Likewise, Gordon discloses allowing a user to choose a movie title from a menu. Gordon at 9:18-28, Fig. 3B (Ex. 1016). After selection, the system determines the title price. Gordon at 9:29-31, Fig. 3B. Thus, Gordon discloses “assessing said speech response” (i.e., assessing the user’s choice of a title) to “create a financial consequence” (i.e., determine the price of the title), as recited in the claim.

414. Because persons of ordinary skill in the art would be motivated to combine Julia with Banker and Gordon (see paragraphs 446-447, below), such combinations also disclose this element.

415. The next element of claim 32 recites:

displaying said financial consequence to create a displayed financial consequence at said user site;

416. The “displaying” step of claim 32 is identical to the “displaying” step

of claim 6. See Paragraph 63, above. Accordingly, Julia discloses this limitation for the same reasons. See Paragraph 355, above. In short, Julia explains that a viewer can use voice commands to select and order pay-per-view movies. Julia at 1:30-43, 10:12-13, 11:31-32. (Ex. 1012). A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase.

417. This limitation is also disclosed in Banker and Gordon. In Banker, the system displays a screen stating “NEXT, PLEASE PRESS [BUY] TO PURCHASE THIS EVENT” to confirm the viewer’s intent to accept the purchase (i.e., the “financial consequence”). Banker at 16:62-17:3, Fig. 6F (Ex. 1015); see Paragraph 355, above. Similarly, after the user selects a movie, Gordon queries the user to see if he wants to pay “a la carte” to view his selection. Gordon at 9:56-63; Fig. 8 (displaying price of movie) (Ex. 1016). See Paragraph 355, above.

418. Because persons of ordinary skill in the art would be motivated to combine Julia with Banker and Gordon (see paragraphs 446-447, below), such combinations also disclose this element.

419. The next element of claim 32 recites:

confirming said displayed financial consequence from said user site to create a financial commitment; and

420. The “confirming” step of claim 32 is identical to the “confirming” step of claim 6. See Paragraph 63, above. Accordingly, Julia discloses this

limitation for the same reasons. See Paragraphs 357-359, above. In short, Julia explains that a viewer can use voice commands to select and order pay-per-view movies. Julia at 1:30-43, 10:12-13, 11:31-32 (Ex. 1012). A person of ordinary skill in the art would know that a pay-per-view movie purchase involves informing the user that they are making a purchase and confirming their intent to do so.

421. This limitation is also disclosed in Banker and Gordon. For example, to purchase a pay-per-view movie in Banker, a user must first press the BUY key, enter a system provided code, and then press the BUY key a second time to confirm the intent to purchase (i.e., the “financial consequence”) to “create a financial commitment” as claimed. Banker at 16:62-17:3, Fig. 6F (Ex. 1015); see Paragraphs 357-359, above. Similarly, after the user selects a movie, Gordon queries the user (i.e., confirming with the user) to see if he wants to pay “a la carte” to view his selection. Gordon at 9:56-63; Fig. 8 (displaying price of movie); see also 6:43-47 (Ex. 1016); Paragraphs 357-359, above.

422. Because persons of ordinary skill in the art would be motivated to combine Julia with Banker and Gordon (see paragraphs 446-447, below), such combinations also disclose this element.

423. The last element of claim 32 recites:

billing said user site based upon said financial commitment.

424. The “billing” step of claim 55 is identical to the “billing” step of claim

19. See Paragraph 63, above. Accordingly, Julia discloses this limitation for the same reasons. See Paragraph 361, above. In short, Julia discloses pay-per-view and video-on-demand functionality. Julia at 1:30-43, 10:12-13, 11:31-32 (Ex. 1012). A person of ordinary skill would also understand that pay-per-view and video-on-demand functionality, such as that provided by the Diva system, would require a purchase, resulting in a bill.

425. In addition, Banker describes a system that allows users to purchase pay-per-view content by pressing the “BUY key” on a television remote and then entering a code provided by the system to begin a purchase sequence. Banker at 16:43-50, 16:62-7:1, Fig. 6F; see also 16:51-55 (Ex. 1015). After following addition steps, the purchase is confirmed and the user is billed. Banker at 17:1-3, 7:60-63, 16:62-17:3. Similarly, Gordon describes a system that allows a user to purchase an on-demand video subscription. Gordon at 9:64-10:9, Figs. 3B and 8 (Ex. 1016). If the user confirms the purchase, the user is billed for the subscription. Gordon at 2:60-63, 3:61-63, 10:39-47. Because persons of ordinary skill in the art would be motivated to combine Julia with Banker or Gordon (see paragraphs 446-447, below), such combinations would also disclose this limitation.

13. Claim 38

426. Claim 38 is a dependent claim that adds the following limitation:

The system of claim 28, wherein the program step responding to said identified speech content from said associated user site is

**comprised, for at least one of said user sites, of the program step of:
responding to said speech content identified as to said user site based upon natural language to create a speech content response of said speech contents identified as to said user site.**

427. Other than minor differences in the preambles (claim 12 recites a “method” whereas claim 38 recites a “system” and adds the language “for at least one of said user sites”), claim 12 is identical to claim 38. See Paragraph 64. Julia discloses this limitation for the reasons discussed in connection with claim 12. See Paragraphs 363-366, above. In short, the remote server 108 in Julia responds to voice requests from user sites (the “speech content identified as to said user site”) by retrieving and sending the requested programming content back to the user site (the “speech content response”). Julia at 2:27-41, 2:61-67, 4:13-17, 6:12-26, 11:60-67 (Ex. 1012). These voice requests are “based on natural language.” Julia at 1:38-42, 1:61-63, 3:37-38, 7:5-7, 7:47-53; see generally Julia at 6:61-12:62 (detailing how the Julia system processes spoken natural language inputs to identify content specified by a user).

14. Claim 39

428. Claim 39 is a dependent claim that adds the following limitation:

The system of claim 28, wherein the program step processing said multiplicity of said received speech channels is further comprised, for at least one of said user sites, of the program step of:

processing said received speech channels from said user site based upon natural language for said user site to create said speech content identified as to said user site.

429. Other than minor differences in the preamble (claim 13 recites a “method” whereas claim 39 recites a “system” and adds the language “for at least one of said user sites”), this limitation is identical to the corresponding limitation recited in claim 13. See Paragraph 65, above. Julia discloses this limitation for the reasons discussed in connection with claim 13. See Paragraphs 368-369, above. In short, Julia discloses using natural language voice requests, such as “Show me the movie ‘Unforgiven’” or “I want to see that movie starring and directed by Clint Eastwood. Can’t remember the title” to request programming content. Julia at 1:38-42, 1:61-63, 10:12-13, 11:31-32 (Ex. 1012). The remote server 108 processes these natural language requests by first performing speech recognition to interpret the words of the user’s request, and then linguistically parsing that information using a natural language interpreter to identify the content requested by the user—the “speech content identified as to said user site” as recited in the claim. Julia at 7:5-7, 7:47-53.

15. Claim 40

430. Claim 40 is a dependent claim that adds the following limitation:

The system of claim 28, wherein said network includes a wired residential broadband network servicing at least part of said

multiplicity of user sites.

431. Julia discloses this limitation. A person of ordinary skill in the art understands a “broadband network” to refer to networks capable of transmitting data at high speeds, such as networks made up of coaxial or fiber-optic cables. This term originated to contrast connectivity means from slower dialup network access using conventional telephone modems. This understanding is also reflected in the ’196 Patent. For example, the ’196 Patent states that “FIG. 2 depicts a typical residential *broadband* network using local loop wiring of the network, as disclosed in the prior art.” ’196 Patent at 4:15-17 (Ex. 1001) (emphasis added). A few lines earlier, the ’196 Patent named Hybrid-Fiber-Copper (HFCop) and hybrid fiber-coaxial (HFC) networks as examples of local loop networks. ’196 Patent at 3:63-4:9. Thus, even according to the ’196 Patent, networks made up of fiber-optic, coaxial, and even copper wires, are broadband networks.

432. Julia is designed to work in “high bandwidth communications infrastructure,” such as “broadband.” Julia at 1:30-32 (Ex. 1012). Julia thus discloses that “[n]etwork 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.” Julia at 4:31-35. The network

106 in Julia is therefore wired (“hardwired connection”) and broadband (coaxial lines, DSL, fiber-optic cable, or traditional copper wire). Moreover, as explained above, the network connects to communications boxes, such as set-top boxes local to the user’s environment (the “user sites”). Julia 3:49-50. The network 106 is thus a “residential” network (because set-top boxes are usually located at user residences) that services “at least part of said multiplicity of user sites” as required by the claim.

16. Claim 41

433. Claim 41 is a dependent claim that adds the following limitation:

The system of claim 40, wherein said wired residential broadband network is a fiber-to-the-curb residential broadband network.

434. Julia discloses this element. As discussed with respect to claim 40, Julia discloses a “wired residential broadband network” that is “residential.” See Paragraph 431-432 above. This limitation additionally requires that the network be “fiber-to-the-curb.” ’196 Patent at 57:33-34 (Ex. 1001). The ’196 Patent does not define the phrase “fiber-to-the-curb”; indeed, the only time this phrase is used is in claim 41 of the patent. Without further guidance, a person of ordinary skill in the art would understand a “fiber-to-the-curb” network to mean a network that runs fiber-optic cables from a central server or headend to a fiber termination node which serves multiple users in the neighborhood of the fiber termination. Other

types of cables, such as coaxial cables or copper wires (twisted pair) connect the multiple user sites to the node. Thus, one example of a fiber-to-the-curb network is a hybrid-fiber-coaxial (HFC) network. In HFC networks, optical fibers run from a central office, or headend, to a neighborhood node. '196 Patent at 3:44-45. "At the node, coaxial cable connects" multiple users to the node. '196 Patent at 3:49-52; see also Quigley at 8:55-60 (defining HFC as networks "wherein optical fiber provides communication to a plurality of fiber nodes and each fiber node typically serves approximately 500 to 2,000 subscribers, which communicate with the node via coaxial cable.") (Ex. 1014).

435. Julia discloses that "[n]etwork 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection." Julia at 4:31-35 (Ex. 1012). A person of ordinary skill in the art reading Julia would understand that one other "type of hardwired connection" that could be used for network 106 is an HFC network. Indeed, the '196 Patent recognized that, at the time of its filing, "[m]ost current cable systems use the popular HFC architecture." '196 Patent at 2:51-54 (Ex. 1001). Other prior art references also demonstrate that HFC networks were common and well known at the time of the '196 Patent. For example, Quigley states that "hybrid fiber coaxial networks are commonly utilized

by cable providers to provide Internet access, cable television, pay-per-view and the like to subscribers.” Quigley at 10:45-48 (Ex. 1014). Therefore, it would have been natural for a person of ordinary skill in the art reading Julia to understand that the network 106 can be an HFC network—a “fiber-to-the-curb” network. Thus, Julia alone discloses this claim.

436. Moreover, both Quigley and Nazarathy also disclose HFC cable networks. For example, Quigley’s system is built for an HFC network. Quigley at 8:55-60, 10:41-11:3 (Ex. 1014). Likewise, the entire disclosure of Nazarathy is directed toward a hybrid fiber coaxial cable network. See, e.g., Nazarathy at Abstract, 9:25-26, (“This invention pertains to broadband HFC distribution systems to digital home terminals”) (Ex. 1013). It would have been natural for a person of ordinary skill in the art to implement Julia in an HFC network as disclosed by Quigley and Nazarathy, given the prevalence of HFC networks used for cable television at the time of the ’196 Patent. Moreover, as discussed below, persons of ordinary skill in the art would have been motivated to combine Julia with Quigley and Nazarathy. See Paragraphs 440-445. The combination of Julia with the HFC networks disclosed in Quigley or Nazarathy thus also disclose claim 41.

17. Claim 42

437. Claim 42 is a dependent claim that adds the following limitation:

The system of claim 40, wherein said wired residential broadband network is a fiber-to-the-home residential broadband network.

438. Julia discloses this limitation. As discussed with respect to claim 40, Julia discloses a “wired residential broadband network” that is “residential.” See Paragraph 431-432 above. Julia also discloses that “[n]etwork 106 is a two-way electronic communications network and may be embodied in electronic communication infrastructure including . . . fiber-optic cable” Julia at 4:31-35 (Ex. 1012). As shown in Figure 1a of Julia, network 106 runs from the remote server 108 all the way to the communications box 104—the user site:

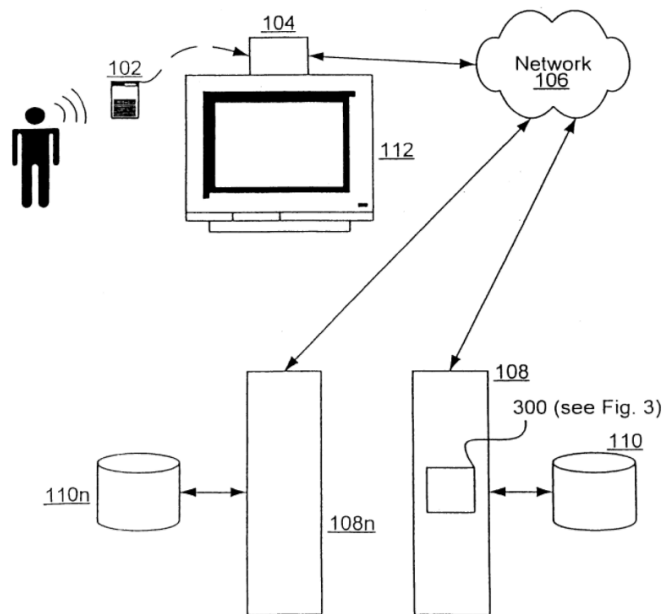


Fig. 1a

439. Thus, Julia discloses a fiber-optic network that runs from a central server (remote server 108) to a user site (communications box 104), as recited by

this claim.

B. A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Julia with the Teaching of Nazarathy and Quigley

440. A person of ordinary skill in the art would have been motivated to combine Julia with Nazarathy and Quigley. Julia, Nazarathy, and Quigley disclose inventions for use in two-way transmission of data over cable networks. For example, Julia, Nazarathy, and Quigley describe transmission of data in cable networks made up of the same types of physical media. Julia states that its network 106 “may be embodied in electronic communication infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.” Julia at 4:31-35 (Ex. 1012). Similarly, Nazarathy states that the cable networks it is designed to function in are “hybrid fiber coaxial cable networks such as used in cable television where two-way digital communications are desired,” with both “coaxial cable and optical fiber portions.” Nazarathy at Abstract (Ex. 1013). Similarly, Quigley states that the cable network it is designed to function in consists of a “hybrid fiber coaxial (HFC) network, wherein optical fiber provides communication to a plurality of fiber nodes,” where a “plurality of subscribers communicate with the fiber node via a common or shared coaxial cable.” Quigley at 8:55-62 (Ex. 1014). In other words, the networked systems of Nazarathy and

Quigley are designed to be used in the very same network infrastructure disclosed by Julia.

441. Further, Julia, Nazarathy, and Quigley describe systems requiring bi-directional communication between a central node (like a headend or a remote server) and multiple, distributed user units. For example, Julia states that its system may be designed to field requests from “simultaneous multiple network users.” Julia at 6:12-27, see also 4:31-35 (“Network 106 is a two-way electronic communications network and may be embodied in electronic communications infrastructure including coaxial (cable television) lines, DSL, fiber-optic cable, traditional copper wire (twisted pair), or any other type of hardwired connection.”) (Ex. 1012). Nazarathy likewise states that its cable modem termination system handle “the upstream digital traffic arriving over digital fibers from multiple nodes into a single digital stream aggregating the individual digital upstream transmissions from the nodes.” Nazarathy at 14:50-54, Figure 9, see also 13:7-9 (“The MAC domain consists of the set of all users served by one or more downstream channels and one or more upstream channels, bound together by the two-way MAC protocol.”) (Ex. 1013). Quigley likewise states that its cable modem termination system “must be capable of facilitating bidirectional communication with any desired one of the plurality of cable modems,” and provides details for how this bidirectional communication may happen. Quigley at

8:46-48, see also 9:18-25 (in the “Detailed Description of the Invention,” noting a “downstream” data rate of up to 56 Mbps and an “upstream” data rate of up to 10 Mbps) (Ex. 1014). Nazarathy and Quigley thus disclose methods for communications that fit directly into the communication needs of Julia.

442. Julia, Nazarathy, and Quigley are also similar in that all three describe transmitting media content in the downstream direction. Julia at 4:10-20, 11:60-67 (Ex. 1012); Nazarathy at 1:15-44, 1:65-2:7 (Ex. 1013); Quigley at 11:4-7, 13:11-21, 16:31-43 (Ex. 1014). Julia and Quigley further describe systems that receive voice commands in the upstream (from user to headend) direction. Julia describes a system that processes natural language audio commands from users and can respond by providing the user with media content, such as video and audio. Julia at 3:35-4:30. Quigley also states that upstream communications can include “voice,” and that its headend, which contains the CMTS, may include “video servers.” Quigley at 11:4, 63:44-46.

443. Nazarathy and Quigley also provide additional details for implementing bi-directional communication systems for cable networks that fit directly into, but are not expressly described, by Julia. For example, Julia states that persons of ordinary skill in the art would appreciate how to implement its system to support receiving and processing multiple simultaneous user requests, but does not provide details regarding how such implementation should occur.

Julia at 6:12-26 (Ex. 1012). Julia also states that its network can be “any type of hardwired connection.” Julia at 4:31-35. Nazarathy and Quigley both describe multiplexing methods that can be used for supporting simultaneous requests, as Julia suggests. Nazarathy describes using time division multiplexing (TDM) and wavelength division multiplexing (WDM) for transmitting signals from multiple users to the cable headend. Nazarathy at Abstract, 13:64-14:46 (Ex. 1013). Nazarathy explicitly teaches that these multiplexed signals are demultiplexed at the cable headend. Nazarathy at 14:62-64 (“Any operations of TDM and/or WDM multiplexing are undone at the HE [cable headend] by corresponding WDM and TDM demultiplexers.”). Nazarathy further teaches that bi-directional cable networks can be implemented as hybrid-fiber-coaxial (HFC) networks. Nazarathy at 1:6-9. Quigley also describes how the “upstream channel” of the cable network can be “partitioned” into a plurality of time intervals. Quigley at 46:32-40 (Ex. 1014). This partitioning is necessary to use “TDMA for upstream communication.” Quigley at 9:56-10:3. Quigley further describes demultiplexing of multiplexed signals. Quigley at 51:49-53. Like Nazarathy, Quigley also explains that HFC networks can be used for cable television networks. Quigley at 8:55-60, 10:41-48. Thus, Nazarathy and Quigley disclose details for multiplexing / demultiplexing signals, which can be implemented in Julia to support receiving and processing simultaneous user requests. A person of ordinary skill in the art,

looking for further details regarding how to implement Julia to support simultaneous requests, would thus naturally consider the teachings of Nazarathy and Quigley because they disclose methods for bi-directional communication that fit directly into the communication needs of Julia.

444. A person of ordinary skill in the art would have recognized the benefit of combining Julia with the teachings of Nazarathy and Quigley. Nazarathy is directed to providing more upstream bandwidth per user by multiplexing signals at intermediate network “hubs” in a hybrid fiber coaxial (HFC) network. Nazarathy at 3:37-43 (Ex. 1013). Quigley is directed to avoiding any undesirable reductions in upstream data rates resulting from a plurality of subscribers competing for bandwidth over a common coaxial cable. Quigley at 1:46-58 (Ex. 1014) In a combined system, the upstream architectures for cable television systems disclosed in Nazarathy and Quigley would be combined with Julia’s speech recognition system, providing more upstream bandwidth per user, allowing for higher fidelity audio commands and a better rate of recognition. The combined system would also improve the user interfaces of the Nazarathy and Quigley systems, as the addition of Julia’s voice control system would allow users to more easily search for, request, and obtain desired programming content. Moreover, a person of ordinary skill in the art would have readily been able to combine the upstream architecture teachings of Nazarathy to Julia’s speech recognition system, to

achieve predictable results, without undue effort or experimentation.

445. Finally, Julia states that persons of ordinary skill in the art would “know” and “appreciate” how to implement the system disclosed in Julia so that it can support simultaneous voice requests from multiple users. Julia at 6:12-26 (Ex. 1012). Because Nazarathy and Quigley describe the same general system architecture as Julia (using fiber optic and coaxial cables for bi-directional communication between multiple network users on one end and a server/headend on the other end), and all three patents relate to sending voice data in the upstream direction, and media content in the downstream direction, a person of ordinary skill in the art would combine the teachings of Nazarathy with Julia as discussed above. This would have been no more than combining prior art elements according to known methods to yield predictable results.

C. A Person of Ordinary Skill in the Art Would Have Been Motivated to Combine Julia with the Teachings of Either Banker or Gordon

446. A person of ordinary skill in the art would have been motivated to combine the teachings of either Banker or Gordon in the Julia system to allow users to order pay-per-view content and movies-on-demand by voice. All three patents (Julia, Banker, and Gordon) relate to user interfaces for control of television systems. Julia at 1:30-38 (discussing “cable, satellite, broadband” and “delivery of movies-on-demand...to the consumer’s home television set”), 3:36-

60, Fig. 1a (showing client display device 112 which may be a television monitor and communications box 104, which may be a set-top box) (Ex. 1012); Banker at Abstract (“Apparatus for providing a user friendly interface to a subscription television terminal comprises a key pad arranged into a plurality of key groupings and an on-screen display controller for generating a plurality of screens for display on an associated television receiver.”) (Ex. 1015); Gordon at 2:41-63 (“The invention is embodied in a combination of software, which provides a so-called “navigator”, and hardware....The graphical user interface functionality is distributed between the service provider equipment and subscriber equipment....Using the inventive equipment, a consumer is capable of further subscribing to packages of on-demand programming, i.e., SOD services, through manipulation of the graphical user interface...As such, through manipulation of the menus, the consumer selects a programming package, becomes a subscriber to that package and is billed accordingly.”); see also 1:8-14 (Ex. 1016). All three patents explicitly discuss pay-per-view functionality in their respective television systems. Julia at 1:30-43 (noting that users can request content, which can include movies-on-demand, and teaches a natural language voice input system to take “full advantage” of that content) (Ex. 1012); Banker at 16:51-55 (“When the terminal of FIG. 3 is in the On mode and tuned to a pay-per-view channel showing an event with its purchase time window open, the user by actuating the BUY key initiates a

buy sequence. Consequently, this key is used to purchase an event.”) (Ex. 1015); Gordon at 1:8-13 (“The present invention relates to an interactive information distribution system such as a video-on-demand (VOD) system.”) (Ex. 1016).

447. Further, a person of ordinary skill in the art would have recognized the benefit of combining Julia with the teachings of either Banker or Gordon that pertain to known features and functionality of pay-per-view systems including, for example, requiring entry of a secret code and confirming user purchases. Julia does not provide any details regarding how pay-per-view can be implemented in its system, other than mentioning the DIVA video-on-demand system. Julia at 6:47-59 (Ex. 1012). Banker and Gordon disclose user interfaces for implementing pay-per-view functionality. Banker at 16:43-50, 16:51-55, 16:62-17:3, Figs. 6F (Ex. 1015); Gordon at 2:60-63, 3:61-63, 8:66-10:9, 10:43-46, Fig. 3B & 8 (Ex. 1016). Persons of ordinary skill in the art would therefore have naturally considered Banker or Gordon in implementing pay-per-view and video-on-demand in Julia. Moreover, the Banker and Gordon approaches reduce the chance of an accidental or unauthorized purchase. It would also have been within the capability of a person of ordinary skill in the art to combine Julia and either Banker or Gordon to permit a user to order pay-per-view movies using the sequence described by Banker or Gordon but using voice commands as in Julia. The combination would have been no more than combining prior art elements according to known methods

to yield predictable results.

X. CONCLUSION

448. For the reasons discussed in this declaration, it is my opinion that the challenged claims of the '196 Patent are unpatentable in view of the prior art.

Acknowledgement of Oath

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1101 of Title 18 of the United States Code.

Respectfully submitted,

A handwritten signature in black ink, appearing to read 'Chris Schmandt', written in a cursive style.

Christopher Schmandt

Dated: December 19, 2017