Nandini Ramakrishnan

Midterm Study Guide

Links:

- https://gatech.instructure.com/courses/327124/files/folder/Readings?preview=43250291
- https://docs.google.com/document/d/1xs9t7QrWsDE7AKR8cfsUKvKQNfduXixnyYSeUwPMckw/edit
-

Accessible Game Design Using Sonification

Universal Design Design that is inherently usable by all prospective users including:

– People with Disabilities

– Children

– Elders

– People without disabilities

Blind Gamer – Ocarina of Time

GDC Talk from EA

• http://www.gdcvault.com/play/1025025/AA

A-Gaming-While

– talk

• http://www.gdcvault.com/play/1025474/AA

A-Gaming-While

– slides

• http://ea.com/able

– Accessibility guides

Multimodal Interaction

• Multiple modes of interaction with a system, facilitating both input and output opportunities

• Redundant but complementary

• Example: A change in health level on HUD also represented by hurt sound

Thief – The Dark Project ('98)

Spatial Simulation

• Situation Awareness

– Environment Description

– Environment Change

– Alerts
• UI controls

Environment Description
• Describe full scene
• Visually: a quick visual scan of the space "look around"
• Auditorily:
– Narration
– Sonification

Environment Change
• Event-based sounds
• Cause and Effect

Alerts
• Command Attention
• Communicate Problem
• Alarm Fatigue

Player Resources
• Inventory
• Avatar state
• Avatar abilities

System Controls
• Graphical UI controls
• Config menus
• Etc.

Player Tasks
• Navigation (find way to desired location)
• Orient (aiming)
• Search and Discovery
• NPC Interactions
• Collaborative Interactions

Sonification
• Presentation of information via non-speech audio

• Informed by intuitive sound concepts (e.g. smaller/further away maps to more quiet)
• Can be informed by existing systems in the environment (e.g. engine sound change in loudness and frequency reflects car change in acceleration and speed)

Sonification
• Mappings:
– Loudness
– Frequency
– Tempo
– Timbre

Sonification – relative change (without running out of head room)
Shepard's Tone

Audio Interfaces for Disabilities
User Needs

- Important for researchers to design technology based on needs assessment from those with vision impairments
- iCare [Panchanathan 2003] – User centric design of assistive devices
    - Researchers not aware of real issues faced by blind and low vision users
    - Researchers ran focus groups to determine issues that are of greatest concern
        - Picking up a book and reading it
        - Knowing who is standing if front of you
        - Web browsing without cognitive overload
            - Screen Readers

Mapping GUIs to Audio

- Was easy with text only displays, speech and Braille pads, typing input
- Not only must the onscreen graphics be presented, but new interaction methods required as well
- Requirements
    - Coherence between graphical and non-graphical apps (don't just make an text only audio app). Needed for collaboration (mental models and synchronization).

    - Efficient Exploration. GUIs can present a lot of information at once in high resolution. Temporal and resolution problems with sound make this hard
    - All graphical information conveyed. Simple things like control type, button states. More difficult icons, multimedia etc.
    - New interaction techniques. Must find equivalent for pointing, clicking, dragging etc.

Mercator

- Mynatt (94), transforms graphical interfaces into audio interfaces
- How can you take a regular GUI app and translate it to audio?
    - GUI is interactive, spatially presented, visually dense
    - Audio interface needs to be efficient, intuitive, and non-intrusive
    - Typically is this accomplished with a "screen reader"
        - Watches the app on the screen and translates what is happening for the user, provides other input methods as well

Mercator

- Mercator takes advantage of "hooks" in Xlib libraries to get information about state changes
    - Hooks exist to provide information to screen readers, testing programs etc.
    - Mercator played role in getting these added to X standard

Mercator: Interface

- Information about the graphical interface is broken down into a tree hierarchy of components (menus, buttons, text areas etc)
    - Fits the mental model that the sighted users have of system, objects in a hierarchy that are used to perform actions
    - Not related to spatial representation on screen (windows occluded, minimized). Some of GUIs are artifacts of the screen size or 2D nature that doesn't need to be translated
    - Based on semantic organization of the applications interface
    - Designed to work with wide variety of applications

Mercator: Interfaces

- Speech and non-speech audio (auditory icons)
    - Edit box, typewriter sound, non-editable text printer sound
    - Toggle button, chain pull light switch
- "Filtears", used to convey attributes of control
    - Grayed out, sound is muffled
- "Hidden" (spatial) information displayed with sound
    - Length of menu mapped to pitch range

- Labels and text are read by the system

Mercator: Navigation

- Can use up,down,left, right keys to navigate the control tree hierarchy.
    - Shortcut keys to jump to points in the hierarchy
    - Also experimented with voice input
- Desktop of applications is just a collection of trees
    - Can move between trees with focus being saved in each
        - Current focus will determine how new information will be presented in the applications
            - Present new info with audio cue and speech
            - Just with audio cue
            - Not at all

Mercator: Meeting the Requirements

- Synchronization
    - Selection of a control in the tree moves the mouse on the screen
- Spatial Exploration
    - Every input moves to new element in tree (no dead space). Breadth first movement in tree exposes main interface components. Depth first, levels of detail

Mercator: Meeting the Requirements

- Conveying Symbolic Information
    - Three levels of auditory cues
        - What is this object?
            - Auditory icon, touching a window sounds like tapping glass, container objects like old wooden box
        - Object's relevant dimensions
            - Length of menu, parameterize pitch of sound
        - Common attributes across components
            - Parameterize the sound, filtears, highlighting and muffling

Mercator: Meeting the Requirements

- Interaction
    - Mouse probably not appropriate
    - Keyboard used
    - Feedback to user about state accomplished through the three types of modes

User Needs

- Pedestrian navigation aids [May 2003] –requirements study of what information is used during urban navigation
    - Landmarks most predominant navigation cue
    - Distance information and street names infrequently used
    - Information also enhances the pedestrian’s confidence and trust.

WayFinding and Navigation

- Popular area of wearable research
- What types of navigation do those with visual impairments need help with?
    - General mobility
    - Crossing the street
    - Reading maps
    - Indoor environments
- What type of interface is appropriate?
    - Audio, speech, haptics, speech, gesture

Wayfinding and Navigation

- CyARM [Ito 2005] – Uses ultrasound to determine how close objects are, transmits information haptically (wire attached to user)
    - Low cognitive load
    - Experiments had participants exploring space and tracking objects

Wayfinding and Navigation

- Personal Guidance System [Loomis 98] – Using audio to guide user along pre-defined route
    - What type of audio output is best?
        - Speech output
            - “point one” (spatialized)
            - “left”
            - “left 80”
        - Spatialized had best results for speed of navigation and distance traveled.

Wayfinding and Navigation

- AudioGPS [Holland 2002] – Non-speech audio conveys distance and heading information to destination or waypoint.
    - Direction
        - Panned tone
        - Tone plus “chase tone” for accuracy
    - Distance
        - “Geiger Counter” paradigm

    - How to minimize obtrusiveness?
        - Use silence, provide volume control

Wayfinding and Navigation

- MoBIC [Petrie 96] – allows user to prepare routes using mapping tool. Routes then used in electronic travel aid. Speech and Braille output.
- Virtual Tactile Maps [Schneider 99]- User can browse map on the go, removes need to memorize.
    - Gesture to explore map, speech/audio output

Wayfinding and Navigation

- “Blind leading the blind” [Gorayska 97] – blind users “draw an acoustic map” to be used pre-trip by other blind users.
    - Binaural recording of route
    - Sharing of acoustic knowledge
    - Leverages blind users existing ability to navigate using acoustic cues

Wayfinding and Navigation

- GuideShoes [Nemirovsky 99] – Provides musical navigational cues in the background.
    - Short musical segments give information on how closely you are following route
    - Aesthetically please
    - Low cognitive load

Wayfinding and Navigation

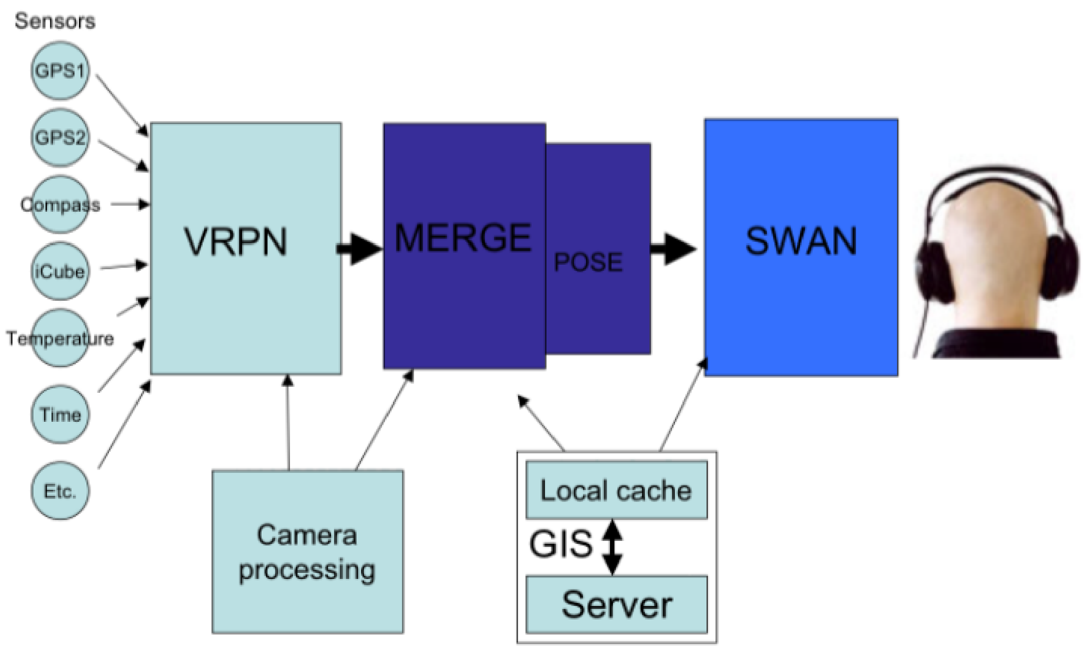
- Interface for navigation [Ross 2000] –Experimented with audio and tactile output for guidance.
    - Help user maintain spatial orientation (e.g. crossing the street)
    - Output types
        - Sonic “carrot” (spatial sound)
        - Speech (e.g. “one o’ clock”)
        - Shoulder tapping
    - Tactile display proved most effective

Wayfinding and Navigation

- SWAN: System for Wearable Audio Navigation (Walker, Dellaert, Wilson)
    - Of use for people who cannot see or cannot look
    - Problem
        - Spatial orientation
        - Veering
        - Learning navigation with late onset blindness

Wayfinding and Navigation

- Many audio interface challenges as well as technical challenges (e.g. sensor fusion, tracking, object recognition, data about environment)
- SWAN system overview



Wayfinding and Navigation

- Types of audio cues (SWAN)
  - Navigation Beacon
  - Object Sounds (furniture, fountains, doorways, etc).
  - Surface Transition (a change in the walking surface)
  - Locations
  - Annotations (additional details about the environment).
- http://sonify.psych.gatech.edu/research/swan/index.html

Wayfinding and Navigation

- Drishti [Ran 2004] – Integrated indoor and outdoor system
  - Input and output via speech
  - User can ask contextual questions about environment.
    - Q: “Where am I?”
    - A: “Living Room”
    - Q: “Sofa?”
    - A: “The sofa is 5 feet away”

Wayfinding and Navigation

- Indoor navigation and object identification [Hub 2004] – “Flashlight” that is used to explore indoor space.
    - Supports exploration of unknown building
    - where you are in a building, how the interior is laid out, and objects around you
    - Keys on device used for input, output via text to speech

Output: Sonification

- Translating data into auditory information
- Large body of work for non-mobile applications
- Useful for those with visual impairments when visual data is translated to auditory output.
    - Most benefit comes when mobile
    - Training usually required to interpret auditory information

Output: Sonification

- Artificial Synesthesia [Foner 99] - Compensate for normal limitations in the human visual system by sonifying portions of visible light spectra that we can not normally see
    - Detect skin cancer
    - See through camouflage

Output: Sonification

- vOICe System, Meijer – Wearable system that translate video from a camera into sound
    - Pixels on the Y axis map to frequencies, their position on the x axis map to the time they will be played
    - Runs on mobile phone

Output: Sonification

- Many systems have broadcast sonar
    - Result that bounces back from obstacles is translated to the user’s audible range
- K-Sonar Cane [Kay 2000]
    - Handheld device attaches to regular cane
- KASPA

Output: Sonification

- Graphs and exact values (Walker)
    - Audio Abacus
    - Midi based toolkit

    - Map pitch range to digits
    - Panning
    - Relative duration can be specified
- Sonification Sandbox
    - Author audio graphs and sonifications
    - Specify mapping of data to dimensions such as timbre, pitch,
    - volume, pan
    - Percussive ticks to indicate scale
    - Repeating tones for min, max, or mean
    - Notifications of data at max, min, mean, changes direction
    - Can display visual version
    - Up to developer to make it sound "good"

Reading

- Braille-recognition system [Mihara 2005] – Mobile
- system utilizing camera, gestural input, and speech
- output
    - Small segment of blind population reads Braille
    - Braille labels in environment (e.g. elevators, room
    - numbers, cash machines)
    - User uses finger to indicate which label to read
    - System reads label aloud

Reading

- Menus
    - Typically text-to-speech is used
        - Incurs downsides of speech
            - Speech
            - Cognitive load
    - Spearcons
        - Speech becomes earcon
        - Very fast speech creates non-speech sound

Accessible Exhibits

- Aquarium (Walker, Weinberg, Balch)
    - Fish type = instrument
    - Movement speed = temp
    - Stereo panning for left/right, filters mapped to vertical
    - At GA aquarium, size of fish = musical pattern
    - Global versus local tracking (i.e. field of view constrained)?

- http://sonify.psych.gatech.edu/research/aquarium/index.html

Customizable Interfaces

- The benefit of having individual devices on the body over devices embedded in the environment
- Difficult to make interface that is accessible to all of the people all of the time
- One solution is to allow user to bring personal interface to target system (cash machine, elevator, home theater etc.)
- V2 - Information Technology Access Interfaces
  - Standards for a Universal Remote Console (URC) framework of components
    - Remote user interfaces and control of devices and services

References

- Keith Edwards, Elizabeth Mynatt,"An Architecture for Transforming Graphical Interfaces", UIST 1994, Nov 2-4
- David A. Ross, Bruce B. Blasch, Wearable interfaces for orientation and wayfinding, November 2000, Proceedings of the fourth international ACM
- conference on Assistive technologies
- Andreas Hub , Joachim Diepstraten , Thomas Ertl, Design and development of an indoor navigation and object identification system for the blind, ACM
- SIGACCESS Conference on Assistive Technologies,2004
- Kiyohide Ito, Makoto Okamoto, Junichi Akita, Tetsuo Ono, Ikuko Gyobu, Tomohito Takagi, Takahiro Hoshi, Yu Mishima, Late breaking results: posters:
- CyARM: an alternative aid device for blind persons, April 2005, CHI '05 extended abstracts on Human factors in computing systems
- Paul Nemirovsky, Glorianna Davenport, Late-breaking results: MIT is the limit: GuideShoes: navigation based on musical patterns, May 1999, CHI '99
- extended abstracts on Human factors in computing systems
- B. Gorayska, J. Marsh, and J. L. Mey, On why the blind leading the blind is a good idea, 'Humanizing the Information Age'. Proceedings., Second
- International Conference on25-28 Aug. 1997 Page(s):173
- Ran, L.; Helal, S.; Moore, S.; Drishti: an integrated indoor/outdoor blind navigation system and service. Pervasive Computing and Communications,
- 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on 14-17 March 2004 Page(s):23 – 30
- Simon Holland, David R. Morse, Henrik Gedenryd, AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface, January 2002, Personal and
- Ubiquitous Computing,Volume 6 Issue 4
- Jochen Schneider and Thomas Strothotte. Virtual Tactile Maps. In J. Ziegler H.-J. Bullinger, editor, 8th International Conference on Human-Computer

- Interaction, HCI International '99, Munich, August 22-27, volume Ergonomics and User Interfaces, Volume 1, pages 531-535, London, 1999. Laurence
- Erlbaum Associates
- Petrie, H., Johnson, V., Strothotte, Th., Raab, A., Fritz, S., Michel, R. MoBIC: Designing a Travel Aid for Blind and Elderly People. The Journal of
- Navigation Vol. 49, No. 1 (1996), 45-52.
- Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (1998). Navigation system for the blind: Auditory display modes and guidance. Presence: Teleoperators
- and Virtual Environments, 7, 193-203.
- Andersen, J.; Seibel, E.;Real-time hazard detection via machine vision for wearable low vision aids,Wearable Computers, 2001. Proceedings. Fifth
- International Symposium on8-9 Oct. 2001 Page(s):182 – 183
- Lee, C.M.; Schroder, K.E.; Seibel, E.J.; Efficient image segmentation of walking hazards using IR illumination in wearable low vision, Wearable
- Computers, 2002. (ISWC 2002). Proceedings. Sixth International Symposium on 7-10 Oct. 2002 Page(s):127 – 128
- Bryant, R., Seibel, E.J., Lee, C.M. and Schroder, K.E. Low-Cost Wearable Low Vision Aid using a Handmade Retinal Light Scanning Microdisplay.
- Journal of Society for Information Display, Special Issue on Selected Papers from the 2004 SID International Symposium.
- Peli E. (2000). Augmented Vision for Central Scotoma and Peripheral field loss. Vision Rehabilitation: Assessment, Intervention and Outcomes.
- Selected papers from Vision '99: International Conference on Low Vision. C. Stuen et al editors, Swets & Zeitlinger, Lisse, p. 70-74.
- Vargas-Martin F, Peli E. (2001) Augmented view for tunnel vision: device testing by patients in real environments. 2001 SID International Symposium,
- Digest of Technical Papers, San Jose, CA. Society for Information Display. P-16, 602-605.
- Stephen Brewster, Lorna M. Brown, Tactons: structured tactile messages for non-visual information display, January 2004, Proceedings of the fifth
- conference on Australasian user interface - Volume 28
- Ivan Poupyrev, Shigeaki Maruyama, Jun Rekimoto, Interaction in the real world: Ambient touch: designing tactile interfaces for handheld devices,
- October 2002 , Proceedings of the 15th annual ACM symposium on User interface software and technology
- Yoshiyuki Mihara, Akihiro Sugimoto, Etsuya Shibayama, Shin Takahashi, Late breaking results: posters: An interactive braille-recognition system for the
- visually impaired based on a portable camera, April 2005, CHI '05 extended abstracts on Human factors in computing systems

- Leonard N. Foner, Artificial synesthesia via sonification: a wearable augmented sensory system, March 1999, Mobile Networks and Applications,
- Volume 4 Issue 1
- Kay, L. Auditory perception of objects by blind persons using bioacoustic high resolution air sonar. JASA, Vol 107, PP 3266-3275, No 6, June 2000

Practice Midterm

1. At the beginning of the semester we discussed the history of audio formats. List 2 analog audio formats, and 2 digital formats
    a. Analog
        i. Phonograph
        ii. DCC- digital compact cassettes
    b. Digital
        i. Voice recorder (vocoder)
        ii. SIGSALY (secure voice communications)
2. Explain the meaning of the terms frequency, period, amplitude, and phase as they relate to a sinusoidal 4Hz sound. Use a diagram of a sound wave to illustrate your descriptions
    a. 1. Frequency:
        i. Frequency refers to the number of complete cycles of a wave that occur in one second. In the case of a 4 Hz sound wave, it completes 4 cycles in one second. The unit of frequency is Hertz (Hz).
    b. 2. Period:
        i. Period is the reciprocal of frequency and represents the time it takes for one complete cycle of the wave to occur. Mathematically, period (T) is the inverse of frequency (f): 1/f . For a 4 Hz sound wave, the period is ¼, or 0.25 seconds.
    c. 3. Amplitude:
        i. Amplitude refers to the maximum displacement of particles in a medium from their rest position when the wave passes through. In the context of sound, it determines the loudness or volume of the sound. A larger amplitude indicates a louder sound, while a smaller amplitude represents a quieter sound.
    d. 4. Phase:
        i. Phase represents the position of a point in time on a waveform cycle. It is often measured in degrees or radians and indicates how far the waveform has shifted horizontally. In simpler terms, it tells you where the wave is in its cycle at a specific moment.
3. Explain what an echo actually is in physical terms. What features of a physical space affect the way audio sounds there and in what ways?

4. A 1Hz sin wave “A” begins playing at time = 0. Then, at time = 0.5 seconds another 1Hz sin wave “B” starts playing too. What is the phase difference in degrees between sin waves “A” and “B”? Graph the resulting sound wave that you would hear starting at time= 0 and extending to time = 2 seconds.
5. Write out the equation (i.e. y(n)=x(n)) for a 4th order filter that produces an output signal that represents a running average of the current and previous input values. In very simple terms what is the effect of this filter? Is this an FIR or IIR filter?
6. You are designing an Analog to Digital Converter (ADC):
    a. How fast would your system need to sample a sound in order to accurately digitize a 12KHz sine wave?
    b. If, instead, the system sampled it at a rate of 18KHz what would the frequency of the reconstructed signal be?
    c. Your boss tells you that you can design the system to use 16 bits per sample or 2 bits per sample for quantization. The 16 bit system will cost a lot more money to manufacture. Which approach do you recommend to your boss? Justify your answer. Draw a diagram to illustrate why your choice is the best.
    d. What is dithering and why would you want to use it in an ADC system?
7. If a sound contains harmonics what does that mean? Draw an example of a sound signal with harmonics.
8. What does a High Pass Filter do? Draw the frequency response of a realistic HPF.
9. List three binaural cues we use to locate sounds in 3D space.
10. Briefly describe three ways that the requirements, use case, design etc. of a wearable computing application are different from that of a smart phone application
11. List 3 reasons why the auditory channel is appropriate for wearable computing applications
12. What is a Head Related Transfer Function used for?
13. What is an advantage of using additive synthesis?
14. Draw an ADSR envelope and label what the letters A,D,S, and R stand for.
15. In AM synthesis you use a “modulator” waveform to control filters that remove frequencies from a “carrier” waveform
16. What is an earcon and what are two advantages of using earcons in an interface?
17. What is an auditory icon and what are two advantages of using auditory icons in an interface?
18. For this filter equation. What is the order of the filter? What are the coefficients (make sure to list all of them)? Is it a Finite Impulse Response (FIR) or Infinite Impulse Response (IIR) filter? y(n) = y(n-1) -5x(n-2) + 2y(n-2)
19. Why does a compact disc use 20 bits to encode 16 bits of data?
    a. A) To account for jitter in playback
    b. B) To reduce the noise inherent to an analog format
    c. C) To avoid certain bit patterns that are hard for the laser to pick up

d. D) To achieve a higher sampling rate via “oversampling”
e. E) To shift quantization noise

20. True or False. If you use 32 bits or more when sampling an analog signal, then there is no quantization noise
21. A pipe organ is a low-tech example of
    a. A) Additive synthesis
    b. B) A generator of white noise
    c. C) Frequency Modulation
    d. D) And auditory illusion based on harmonics
    e. E) A digital sound format
22. A white noise signal with bandwidth between 20Hz and 20KHz is passed through a filter. The resulting signal has a bandwidth that spans 20 Hz to 1KHz. The type of filter that was used is a:
    a. A) High Pass Filter
    b. B) Low Pass Filter
    c. C) Band Pass Filter
    d. D) Comb Filter
    e. E) An Impulse
23. A digital time domain representation of a sound signal consists of:
    a. A) The parameters needed to control a filter bank that will resynthesize the analog signal
    b. B) Instantaneous frequency values for a signal over very small windows of time
    c. C) A set of coefficients that characterize the behavior and complexity of a filter
    d. D) An array of “bins” that represent the power present in the signal for different frequency ranges
    e. E) Discrete values representing signal intensity that were collected at a defined sampling rate
24. To remove a sound stream from the environment, noise cancelling headphones do the following:
    a. A) Analyze the original signal and synthesize the same signal Π radians out of phase and play it back.
    b. B) Perform an FFT on the signal and then construct a filter bank of low and high pass filters, which are then applied.
    c. C) Apply dithering
    d. D) They are built using “planar magnetic” drivers for producing sound
    e. E) Make use of psychoacoustic research to generate masking tones that “conceal” unwanted frequencies.
    f. F) Plug the pinna so the listener cannot locate, and thus consciously notice, the unwanted sound
25. True or False. This is a finite impulse response filter, $y(n) = 3x(n-2) -6\,x(n) + 8x(n-5)$?

26. The coefficients of a filter (like the one shown in question #25) tell us the following about a filter:
    a. (A) The order of the filter
    b. (B) Whether it is an FIR or IIR filter
    c. (C) The behavior of the filter (i.e. how it will effect the input signal)
    d. (D) All of the above
    e. (E) None of the above
27. What do you do at the beginning of the analog-to-digital process?
    a. A) Start converting the continuous signal into discrete sample values
    b. B) Convolve the signal with the impulse response
    c. C) Create a buffer of random sample values
    d. D) Do an FFT on the signal to determine what frequency range is present and then change the sampling rate
    e. E) Apply a low pass filter so as to avoid aliasing
28. True or False. A "Mondegreen" is an example of how our sound perception can be affected by contrasting visual and auditory cues.
29. 29. (2pt) The ear structure that helps us resolve ambiguities when locating sounds in 3D space is:
    a. A) Eardrum
    b. B) Pinna
    c. C) Cochlea
    d. D) Ossicles
    e. E) Cilia
30. 30. (2pts) The following is an example of using audification for sonifying ocean tide:
    a. A) Playing the raw data set as though they are sound samples
    b. B) Generating a continuous wave sound and then modifying the pitch of the sound based on the data values
    c. C) Picking a base musical segment that is mapped to a range of data and then changing the volume of the music based on the actual data value
    d. D) Moving a basic sin wave tone around in 3D space, where the data values determine the x,y,z location of the sound
    e. E) Using the data set as input into a physical simulation of the ocean that then synthesizes realistic sounds.
31. True or False? Nomadic Radio was a system used to study how anesthesiologists could use sound cues to monitor a patient's vital signs during surgery.
32. True of False? The Arkola experiment studied how people used sound cues to collaboratively control a virtual soda bottling plant
33. You are designing a computer system for the military that will help the sailors on a submarine during emergency situations to diagnose and fix problems. You are trying to

convince the project leader, who doesn't want to bother with audio output as part of the system, why you all should be including audio in this environment.

a. List three reasons that using an auditory interface in this application and environment will be challenging.
b. List three reasons to use audio that you might include in your argument with your ignorant boss.